Universidad Carlos III de Madrid

TESIS DOCTORAL

Título de la tesis:
Three Essays on Financial Markets

Autor: Iván Blanco

Director: Alejandro Balbás

Doctorado en Economía de la Empresa y Métodos Cuantitativos

Departamento de Economía de la Empresa

Getafe (España), Marzo de 2016

TESIS DOCTORAL

TÍTULO DE LA TESIS:

Three Essays on Financial Markets

Autor: Iván Blanco

Director: Alejandro Balbás

Firma del Tribunal Calificador:

	Nombre y Apellidos	Firma
Presidente		
Vocal		
Secretario		
Suplente		

Calificación:

Getafe, de de 2016

To my family

Contents

List of Figures

List of Tables

Acknowledgements

"Nobody ever figures out what life is all about, and it doesn't matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough."

Richard Feynman.

Firstly, I would like to dedicate my thesis to my family. A very special thanks goes out to my wife Mercedes, who has suffered all of the ups and downs of my Ph.D with me. Mercedes, you have enjoyed and growth with me in this life project and you have given me the biggest gift, my daughter Noa. I would also like to thank to my mother, grandmother and the rest of my family. Thanks a lot.

I am immensely grateful to my supervisor, Alejandro Balbás, for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge.

I place on record, my sincere thank you to all my coathors: David Werheim, Rosa Rodriguez, Ignacio Peña, Sergio García and José Garrido. I have been extremely lucky to have the opportunity to work with all of you. Besides, I also take this opportunity to express gratitude to the Universidad Carlos III de Madrid for providing the funding which allowed me to undertake this thesis and the entire faculty members for providing me the support and friendship throughout my time at the University.

Last but not least, I am very grateful to all my friends and classmates for supporting me spiritually throughout writing this thesis and my life in general.

Chapter 1

The Bright Side of Financial Derivatives: Options Trading and Firm Innovation

1.1 Introduction

Innovation is the main driver of growth and the wealth of nations. As emphasized by Porter (1992, p. 65), *"[t]o compete effectively in international markets, a nation's businesses must continuously innovate and upgrade their competitive advantages. Innovation and upgrading come from sustained investment in physical as well as intangible assets."* Given the importance of innovation for competitiveness, it is a priority to understand those factors that determine incentives to innovate at the firm level. There has been much debate on the role of financial markets in promoting innovation. While developed capital markets can improve the efficiency of long-term resource allocation through their monitoring and disciplining mechanisms, the need to meet quarterly or annual financial objectives gives rise to adverse externalities that may impair firms' incentives to innovate (Holmström, 1989; Porter, 1992).[1]

In this paper, we focus on one cornerstone of public equity markets, namely financial derivatives. Specifically, we study whether the volume of equity options written on the underlying asset encourages or impedes firm innovation. Since the beginning of the new century, the total equity options volume traded on U.S. exchanges has grown exponentially, from 676 million contracts in 2000 to over 3,727 million contracts in 2015.[2] Unlike stock market listings, where firms apply, options listings are exogenous to firm decisions; they are made within exchanges. These exchanges are self-regulating institutions that are members of the Options Clearing Organization (OCC) which operates under the jurisdiction of the Securities and Exchange Commission (SEC) (for exchange-listed options). Because the SEC plays an important role in determining the eligibility criteria for securities in options trading, this topic is of particular interest to policy makers.[3]

Did the significant rise in the volume of trading undermine innovative efforts or did it encourage firms to invest in innovation? We argue that for firms that are listed on options markets, greater trading activity is associated with an increased propensity to innovate. The literature suggests that active options markets alter incentives for market participants to gather private information that is especially relevant for long-term investments, and trading on such information makes stock prices more efficient (e.g., Cao, 1999; Chakravarty, Gulen, and Mayhew, 2004; Pan and Poteshman, 2006; Hu, 2014). If

[1]Laurence D. Fink, chairman and CEO of BlackRock, recently summed this up in a letter to S&P 500 CEOs that BlackRock invests in (Business Insider, April 14, 2015): *"Over the past several years at BlackRock, we have engaged extensively with companies, clients, regulators and others on the importance of taking a long-term approach to creating value. We have done so in response to the acute pressure, growing with every quarter, for companies to meet short-term financial goals at the expense of building long-term value. This pressure originates from a number of sources – the proliferation of activist shareholders seeking immediate returns, the ever-increasing velocity of capital, a media landscape defined by the 24/7 news cycle and a shrinking attention span, and public policy that fails to encourage truly long-term investment."*

[2]See Options Clearing Corporation.

[3]See Mayhew and Mihov (2004) for initial listing requirements.

stock prices are more efficient, other types of (perhaps less-informed) investors learn more about the fundamental value of the firm, which reduces some of the asymmetric information problems connected to R&D. Because prices play an active role (i.e., managers learn from prices) when investment decisions are made, this should then provide firm management with more incentives to engage in value-enhancing innovative activities. Note that informed agents in financial markets can ameliorate asymmetric information related to innovative activities is widely recognized in the literature (e.g., Hall and Lerner, 2010; Aghion, Van Reenen, and Zingales, 2013; He and Tian, 2013).[4]

In this paper, we focus on whether options trading spurs firm innovation in the context of R&D-intensive industries. We believe that these firms provide an ideal research setting for our study. For firms that invest more heavily in R&D, innovation is a core component of their competitive strategy, but they might also be forced to make only partial disclosure and be subject to a larger degree of information asymmetry (Bhattacharya and Ritter, 1983; Anton and Yao, 2002). It follows that these firms are more likely to be undervalued by equity holders and have a greater exposure to hostile takeovers (Stein, 1988). Moreover, survey evidence by Graham, Harvey, and Rajgopal (2005) shows that managers in technology-intensive industries are more prone to sacrifice long-term sustainability to meet desired short-term earnings targets, relative to managers in other industries, due to their personal wealth and career concerns. They explain that meeting earnings benchmarks (particularly the same quarter earnings from the previous year) helps to maintain a firm's current stock price. Taken together, if the enhanced informational efficiency induced by options leads to better monitoring by reducing information asymmetries, making firms more willing to invest in innovation, we claim that this mechanism is particularly relevant for firms operating in R&D-intensive industries.

To test this conjecture, we assemble a rich and original dataset containing time-varying information on standard measures of innovation based on U.S. patent data, R&D, options trading, governance, etc. To approximate the total annual dollar options volume, we use the approach proposed by Roll, Schwartz, and Subrahmanyam (2009). We run panel data regressions on a sample of 548 publicly traded U.S. firms during the period from 1996 to 2004. This sample consists of large firms that are active in five broadly defined high-tech sectors, where we observe high patenting propensities, and patents have been recognized as a meaningful indicator of innovation at the firm level (as explained in Section 1.3).

Our baseline test reveals a positive association between innovation and options trading. Options trading has a positive impact on R&D spending but a larger positive effect

[4] If we believe that informed agents can reduce information asymmetries related to innovative activities and that the stock market is an efficient resource allocation mechanism, then the "prospective role" (Dow and Gorton, 1997) in which stock prices provide managers with relevant information for investment decisions could generate the same prediction. Our focus on the disciplining role of stock prices (as in Holmström and Tirole, 1993) is a natural choice for understanding the role of options trading in innovation, although we consider the two approaches to be complementary.

on the quality and/or productivity of R&D (i.e., citations per dollar of R&D invested). These results are robust to using alternative sub-samples, alternative measures of innovation, the inclusion of a wide range of control variables, lagged explanatory variables and several econometric models. While these findings are consistent with the beneficial effect of the production and aggregation of information in options markets, we have concerns that our results could be biased if informed agents trade on the basis of unobservable characteristics that are correlated with options volume and innovation. We account for such selection issues by weighting sample observations using their propensity score of having high levels of options trading and by estimating two-stage least squares (2SLS) models using moneyness and open interest as instrumental variables. Overall, our identification tests suggest that the positive correlation between options trading and innovation is not simply driven by self-selection.

We extend these baseline results in two main directions. First, we examine the link between options trading and three measures of innovative direction: (i) a measure based on the diversity of patents applied for by the firm across technological classes, (ii) the Hall, Jaffe, and Trajtenberg (2001) measure of patent originality and (iii) a measure of risk-taking behaviour based on the standard deviation of citations received across patents. The results suggest that more active option markets are associated with a change in direction and not just an increase in R&D spending and productivity.

Second, we attempt to identify the underlying economic mechanism through which this link occurs. Our results could be explained by two hypotheses. On the one hand, the results could be driven by the reasoning that poorly governed managers prefer to avoid the difficult decisions and costly efforts associated with innovation and that the information conveyed by more active options markets "forces" managers to innovate if they are a priori reluctant to do so (i.e., managers prefer the quiet life as in Bertrand and Mullainathan (2003)). On the other hand, the results could be consistent with the prediction that increased monitoring "shields" managers against those reputational consequences (i.e., career concerns as in Holmström (1989, 1999)) that are more likely to occur when managers invest in innovation. Potential consequences occur because innovation involves a high probability of failure, and the innovation process is unpredictable and idiosyncratic, with many future contingencies that are impossible to foresee. In line with recent evidence by Aghion, Van Reenen, and Zingales (2013) in the context of institutional investors, we find strong support for the career concerns story. We show that the positive effect of options trading on innovation is more pronounced when product market competition is more intense, when CEOs are less "entrenched", and for younger CEOs. Moreover, we provide evidence that the positive effect of more active options markets on innovation is magnified for firms that face a decline in profitability and remains substantial even after accounting for executive compensation schemes.

Although we follow standard procedures in using patent counts weighted by forward citations as a proxy for innovation, we must admit that one of the main limitations of our study is that we cannot completely exclude the possibility that our results may be partially driven by managerial signalling motives. This is because one common downside in studies based on patents is that they are an indirect measure of innovation and contain no information on non-patentable inventions or inventions held in secrecy. We believe, however, that limiting our study to industries in which patenting represents the most important mechanism used by these firms to protect their intellectual property for appropriability and/or strategic reasons mitigates such concerns.

While there is a growing literature that links a variety of financial market characteristics to innovation, to the best of our knowledge, such an analysis of the relationship between options trading and innovation has not previously been undertaken. Empirical studies have examined, for instance, the effect of institutional ownership on innovation (Aghion, Van Reenen, and Zingales, 2013), analyst coverage (He and Tian, 2013), credit supply (Amore, Schneider, and Žaldokas, 2013), stock liquidity (Fang, Tian, and Tice, 2014), leveraged buyouts (Lerner, Sorensen, and Strömberg, 2011), investors' failure tolerance (Tian and Wang, 2014), the decision to go public (Bernstein, 2015) and the development stage of financial markets (Hsu, Tian, and Xu, 2014). There is very little on the role played by options (or more general financial derivatives) in the R&D process of publicly traded firms.

However, the possibility that active options markets are beneficial to the firm has also been examined by another paper. Specifically, Roll, Schwartz, and Subrahmanyam (2009) find that options trading activity increases firm value through its impact on price informativeness. However, because greater informational efficiency tends to make an asset more valuable because it reduces the risk of investing in it, these results require further examination. Although several other studies also conclude that resources are allocated more efficiently if prices convey more information, which in turn leads to greater firm value (e.g., Khanna, Slezak, and Bradley, 1994; Dow and Gorton, 1997; Subrahmanyam and Titman, 1999; Durnev, Morck, and Yeung, 2004; Chen, Goldstein, and Jiang, 2007), there is little empirical evidence of this effect on innovation. We view our study as complementary to Roll, Schwartz, and Subrahmanyam (2009) because we take option markets' effect on prices as given and aim to explain how this influences firms' incentives to innovate. Thus, the main contribution of our paper is to provide a direct link between options trading and the extent to which the firm allocates resources to innovation.

The remainder of the paper is organized as follows. Section 1.2 discusses the related literature in greater detail. Section 1.3 describes the sample, the measurement of variables and descriptive statistics. In Section 1.4, we present our main results. In Section 1.5, we discuss the underlying mechanism through which options trading may affect innovation. Section 1.6 concludes the paper.

1.2 Related Literature

Our paper borrows from different strands of the literature. Our starting point is the recognition that options stimulate informed trades and that the informational benefit of options depends on the trading volume. Almost 40 years ago, Ross (1976) was the first to argue that options trading can convey important information in a market with information asymmetry by expanding the contingencies that are covered by traded securities. Apart from reducing information asymmetry, Black (1975) notes that informed traders could use options markets as an alternative venue for trading because option contracts provide higher leverage. Easley, O'Hara, and Srinivas (1998) argue that options can be more attractive for informed traders because the availability of multiple contracts confronts uninformed traders with substantial challenges. In a similar vein, Cao (1999) suggests that agents with private information should be able to trade more effectively on their information in the presence of options, thus improving price informativeness. Moreover, options are a mechanism for trading on information about future equity volatility, which allows investors with information about stock price volatility to benefit from options (Ni, Pan, and Poteshman, 2008). These notions are further supported by Chakravarty, Gulen, and Mayhew (2004) and Pan and Poteshman (2006), who find that options order flows contain information about the future direction of the underlying asset. More recently, Hu (2014) shows that an options-induced imbalance significantly predicts future stock returns. Taken together, these works provide strong support for the conjecture that informational efficiency may be greater in the presence of options.

A firm's informational benefit from options, however, should depend on the volume of options traded, beyond the presence of an options market on the firm's stock per se (as in Roll, Schwartz, and Subrahmanyam, 2009). For example, due to the maxim that "liquidity attracts liquidity", informed agents would be more willing to trade on their private information in markets with high trading volume since they allow them to camouflage their trades (Kyle, 1985; Glosten and Milgrom, 1985). In contrast, if informed traders perceive a low-liquidity options market, they optimally desist from trading and this belief becomes self-fulfilling (Admati and Pfleiderer, 1988; Chowdhry and Nanda, 1991). It follows that the enhancement of the benefit from listing should be directly related to whether the market for the listed options has sufficient volume because then informed traders would be more active.

Second, our paper builds on the literature that interacts information production (i.e., price informativeness) with investment decisions in firms. The idea that the production and aggregation of information as a consequence of trading between speculators and investors can be useful for the provision of incentives in firms is a relatively recent one. Specifically, Holmström and Tirole (1993) and Faure-Grimaud and Gromb (2004) examine the role of price informativeness in disciplining managers and providing incentives to insiders to engage in value-increasing activities. Dow and Gorton (1997) show that, in equilibrium, the information contained in stock prices can be used to guide investment decisions because managers are compensated based on future stock prices. Subrahmanyam and Titman (1999) study a setting in which investors may obtain information unavailable to firm insiders that is useful in making investment decisions. They show that if such information is freely available to outsiders, the firm chooses to go public. Empirically, for example, Durnev, Morck, and Yeung (2004) show that U.S. industries and firms exhibiting larger firm-specific return variation make better capital budgeting decisions. The findings in Chen, Goldstein, and Jiang (2007) suggest that firm managers learn from private information concerning their own firms' fundamentals contained in stock prices by incorporating stock price information into corporate investment decisions. Foucault and Gehrig (2008) show that cross-listing enables firms to obtain more precise information on the value of their growth opportunities, which allows managers to make better investment decisions. Finally, Ferreira, Ferreira, and Raposo (2011) provide evidence that if prices are more efficient, the stock market is able to play a monitoring role that can reinforce internal and external monitoring mechanisms, although the sign of this relationship is ambiguous (i.e., they can interact as either complements or substitutes).

Our particular focus is on how the enhanced informational efficiency induced by options can "boost" managerial incentives to invest in innovation. Stein (1989) shows that even in rational capital markets, firms take actions to improve current earnings at the expense of lower future earnings in an attempt to misguide the market. Shleifer and Vishny (1990) offer a different argument that leads to the same conclusion. Because arbitrage is cheaper for short-term assets than for long-term assets, the latter must be more mispriced in equilibrium for net returns to be equal. It follows that managers may forgo investment opportunities in long-term projects because the uncertainty of these assets can take a long time to disappear. The empirical literature has shown evidence consistent with managerial short-termism in publicly traded firms. For example, Asker, Farre-Mensa, and Ljungqvist (2015) find that compared to unlisted firms, listed firms tend to invest less and their investment levels are less sensitive to changes in investment opportunities. Bushee (1998) shows that firms are more prone to cut R&D in response to a decline in earnings when a very large proportion of institutional owners are investors that often trade in and out of individual stocks.

Based on the streams of literature reviewed above, we argue that a potential solution to the distortion of innovative investment due to agency problems is active options markets. The intuition is the following. In the presence of option market participants who engage in monitoring, informed agents move the stock price towards the fundamental value and thus cause it to more closely reflect the effort exerted by the manager to enhance long-term risky investment decisions. Because other financial market participants (especially firm investors) may have difficulties properly evaluating managerial investment decisions in innovation (Stein, 1988), they can use stock prices as a signal of whether informed traders agree or disagree with the allocation of corporate resources and can decide whether to take action (as in Edmans and Manso, 2011). For example, if investors discover that the manager is good despite bad public information, they will be more willing to retain their shares because they will expect higher returns. Alternatively, they can directly use the threat of disciplinary trading and can sell more if information turns out to be negative, causing the stock price to drop.

1.3 Data and Methods

We examine the effect of options trading on innovation in the context of publicly traded U.S. firms in the following five industries: (i) pharmaceuticals (SIC code 283), (ii) industrial and commercial machinery and computer equipment (35), (iii) electronics and communications (36), (iv) transportation equipment (37), and (v) instruments and related products (38). A trade-off made in the design of our study was to limit the sample to these five industries and not to consider the entire manufacturing universe. We ensured that these industries represent a broad spectrum.[5] Nevertheless, we exercised caution in selecting these specific industries for several reasons. First, R&D has been and continues to be vital for the long-run competitive advantage of firms operating in these industries. In fact, these sectors have the highest ratio between R&D expenditure and net sales among all industries (OECD, 2013). Second, these industries also form an apt context because of how they protect and document their inventions. Patenting (on which our dependent variables are based) is an important mechanism to protect intellectual property (Levin, Klevorick, Nelson, Winter, Gilbert, and Griliches, 1987) and firms tend to patent most patentable inventions. In particular, Mansfield (1986) shows that our sample industries are characterized by high patenting propensities relative to most other industries. Third, patents are a meaningful measure of innovation in these industries. The association between patents and technological innovation is likely to be stronger in industries in which patents provide firms with fairly strong protection of their intellectual property.[6] Acs, Anselin, and Varga (2002) conclude that the measure of patented inventions provides a fairly good, although not perfect, representation of innovative activity in these five industries. Therefore, patents have been extensively used in earlier research to understand the innovation processes of firms within these industries (e.g., Katila and Shane, 2005; Coad and Rao, 2008; Rothaermel and Alexandre, 2009).

[5]In 2004, these industries collectively included approximately 35% of all publicly traded U.S. manufacturing firms drawn from the Compustat database.

[6]Strictly speaking, patents are inventions. As Freeman and Soete (1997, p. 22) notes, they represent *"[...] an idea, a sketch or a model for a new improved device, product, process or system. Such inventions may often (not always) be patented but they do not necessarily lead to technical innovations."*

We use firm-level data on innovation and options trading from several data sources. Our starting point is the Compustat universe which contains detailed information for all U.S. publicly listed firms since the mid-1950s. We identified all firms traded on NYSE, AMEX or NASDAQ with accounting data available between 1996 and 2004. To mitigate backfilling bias, we require firms to be listed on Compustat for three years before including them in the sample. Our main Compustat items are sales (SALE); a capital-labour ratio constructed from the net stock of property, plant, and equipment (PPENT) and the number of employees (EMP); and R&D expenditure (XRD). R&D is used to create R&D capital stocks calculated using a perpetual inventory method with a 15% depreciation rate following the method described in Hall, Jaffe, and Trajtenberg (2005).

Firm-level patent data are obtained from the latest version of the NBER patent database which contains approximately three million patents granted by the United States Patent and Trademark Office (USPTO) and citation information from 1976 to the end of 2006 (Hall, Jaffe, and Trajtenberg, 2001; Jaffe and Trajtenberg, 2002).[7] We use patents that are ultimately granted, dated by the year of application, which approximates the year when the invention is completed because the patent system provides incentives to file quickly. To match Compustat firms with U.S. patent assignee codes, we begin with the name-matching tool of Bessen (2009) and then search by hand for variations on the names in our panel. Our sample ends in 2004 because many patent applications filed in the later years (i.e., 2005 and 2006) might still be under revision (Hall, Jaffe, and Trajtenberg, 2001).

For data on options trading, we use OptionMetrics. This database contains information on the daily number of contracts traded for each individual put and call option on U.S. publicly listed equities along with daily closing bid and ask prices from 1996 onwards. The sample is selected to include firms with positive options volume to maintain comparability, as firms without options listings tend to be small (Mayhew and Mihov, 2004) with different structural relationships between innovation and the right-hand variables.[8] To approximate the total annual dollar options volume, we use the approach in Roll, Schwartz, and Subrahmanyam (2009). Specifically, for each stock, we first multiply the total trade in each option by the end-of-day quote midpoint for that option and then aggregate this number annually across all trading days and all options listed on the stock.

To calculate other control variables and the variables used for exploring underlying mechanisms, we collect institutional ownership information from Thomson Reuters' CDA/Spectrum Institutional Holdings dataset (form 13F), corporate governance information from the RiskMetrics database, analyst coverage data from the Institutional Brokers'

[7] See NBER Patent Data Project website.

[8] For example, Acs and Audretsch (1988) show that small firms spend disproportionately less on R&D, but they appear to benefit more from R&D investments, suggesting they are more efficient at R&D than their larger counterparts.

Estimate System (I/B/E/S) database, CEO age and compensation from ExecuComp, stock price information from the Center for Research in Security Prices (CRSP), intraday trades and quotes for constructing stock illiquidity measures from the Trade and Quote (TAQ) database, and information on each firm's alliances and joint ventures from the Securities Data Company (SDC) Platinum database.

These datasets do not overlap perfectly; thus, our baseline regressions run from 1996, the first year options trading data are available, to 2004, the last year when we can realistically construct innovation measures based on patent data. Although the exact number of observations depends on the specific regression, the baseline sample for which we estimate the equations contains 3,271 observations on 548 firms.[9]

1.3.1 Dependent variables

Our primary measure of innovation is a *future* citation-weighted count of U.S. patents. We prefer patents weighted by citations as an indicator of innovative "output" over simple counts because patent citations can better reflect the technological and economic "importance" or "value" of the underlying invention (Trajtenberg, 1990; Albert, Avery, Narin, and McAllister, 1991). Specifically, the use of patent citations exploits the fact that patent applications must acknowledge "prior art", in which light they need to meet the requirements for patentability, i.e., U.S. patent law requires an invention to be novel, non-trivial and susceptible to industrial application for a patent to be granted.[10] These citations serve an important legal function because they can delimit the scope of the property rights awarded to the inventor. U.S. patent applicants are legally required to disclose any knowledge upon which their inventions are based. This prior art is typically referenced through citations provided by patent applicants (inventors or their lawyers) and patent examiners. Because of this important legal function, the economics of innovation literature has frequently used the number of forward citations received by a patent as an indirect measure of its value (e.g., Pakes and Griliches, 1980; Harhoff, Narin, Scherer, and Vopel, 1999; Aghion, Bloom, Blundell, Griffith, and Howitt, 2005; Hall, Jaffe, and Trajtenberg, 2005; Aghion, Van Reenen, and Zingales, 2013). To control for the fact that citation counts are inherently truncated, we employ three strategies. First, we estimate until 2004, allowing for a two-year window of forward citations for the last cohort of patents in the data. Second, we include a full set of time dummies, which accounts for the fact that patents taken out later in the panel have less time to be cited than patents taken

[9]Our sample faces another restriction from the overall Compustat database. Because our preferred regressions use firm fixed effects, we condition our sample on firms that received at least one citation and had at least two years of non-missing data for all variables between 1996 and 2004. Thus, we drop firms from the Compustat/USPTO match that patented prior to 1996 but not in the 1996–2004 period, and of those that did patent, we drop those that did not receive citations.

[10]See 35 U.S.C. 102 Conditions for patentability; novelty.

out earlier in the panel. Third, we also perform our estimations using simple unweighted patent counts.[11]

We consider several additional innovation metrics. First, we use R&D expenditure as a measure of innovation inputs. Because more than 50% of firms in the entire Compustat database do not report R&D expenditures, we follow common practice in the literature by replacing missing values with zeros, although we obtain similar results when we drop these observations or interpolate over any gaps of three years or less.[12] Second, given that self-citations may differ from other citations in various ways (Hall, Jaffe, and Trajtenberg, 2005), we weight patents by the number of non-self forward citations. Finally, in a series of extensions, we examine changes in the direction of innovative efforts. To proxy for the direction of a firm's activities in its innovative process, we use the diversity of activities (i.e., the dispersion of the firm's patent portfolio across technological classes), originality-weighted patent counts (i.e., the dispersion of backward citations across technological areas) and a measure of risk-taking (i.e., the standard deviation of forward citations across a firm's patents).

It is important to note, however, that using patent data to measure innovation also has limitations. In particular, not all firms patent their inventions because some inventions do not meet the patentability criteria, and others are not patented for strategic reasons. Moreover, firms differ in their patenting propensity, and the degree to which these factors are problematic varies substantially across industries (e.g., Levin, Klevorick, Nelson, Winter, Gilbert, and Griliches, 1987; Cohen and Levin, 1989; Griliches, 1990). We believe that limiting our study to specific industries in which patents are a meaningful indicator of technological activities reduces such concerns because other factors that may affect patent propensity are relatively stable within such a context (Cohen and Levin, 1989; Griliches, 1990). Because firms may differ in their patenting propensity for unobserved reasons even in R&D-intensive industries, we treated this problem as one of unobserved heterogeneity across industries and firms, and control for such variations in our statistical analysis.[13]

[11]We also experimented with adjusted citations, taking into account systematic differences in the number of citations each patent receives across application year and technological class (Hall, Jaffe, and Trajtenberg, 2001). This delivers very similar results to the unadjusted citations results presented here.

[12]Note that the fact that a firm does not report R&D expenditures in its financial statement does not necessarily imply that the firm is not engaging in R&D. Because this information is public, a firm could decline to report for strategic reasons.

[13]For example, one concern might be that our analysis includes firms in "complex" (i.e., SIC codes 35, 36, 37 and 38) and "discrete" industries (i.e., pharmaceuticals; SIC code 283) in the sense proposed by Cohen, Nelson, and Walsh (2000). The authors define complex (discrete) industries as those in which a given technology is protected by many (few) patents. One might therefore observe a lower number of patents generated by firms in discrete industries, but this does not necessarily imply that these firms are less innovative. Moreover, one may argue that our industry classification is too broad to isolate, for example, the role of highly innovative biotechnology companies within the pharmaceutical industry. We account for this potential bias in our regressions by using the most detailed industry classification available (i.e., four-digit SIC code).

1.3.2 Descriptive statistics

Table 1.1 provides summary statistics of the main variables used in this study.[14] Our sample firms are large: $494 million in net sales at the median and 2,400 employees. On average, a firm in our sample has 62 granted patents per year and subsequently receives 294 citations for its patents, which is comparable to previous studies (e.g., Aghion, Van Reenen, and Zingales, 2013). The citation series is highly skewed, with a median of 15. Due to the right-skewed distribution of cite-weighted patents, we use the natural logarithm as the main innovation measure in our analysis. To avoid losing firm-year observations, we add one to the actual values when calculating the natural logarithm. The options volume measure has a mean value of $157 million and a median value of $700 million. Regarding the other variables, an average firm invests $287 million in R&D, approximately 51% of shareholders are institutional investors, its return on assets is 9%, and 22.5 years have passed since its inclusion in Compustat.

Table 1.1: Descriptive statistics

	Mean	StdDev	Min	Median	Max	Observations	Source
Cite-weighted patents	293.7	1181.3	0	15	18950	3271	USPTO
Patents	62.4	185.2	0	7	2355	3271	USPTO
Non-self cite-weighted patents	233.4	906.04	0	12	17188	3271	USPTO
Originality-weighted patents	30.8	89.3	0	4.0	1158.3	3245	USPTO
Std. Dev. of patent citations	4.9	6.02	0	2.8	66.5	2382	USPTO
Innovative diversity	0.62	0.21	0	0.67	0.94	1526	USPTO
Options volume (in $m)	156.7	700.3	0.00018	8.5	15134.5	3271	OptionMetrics
Moneyness	0.29	0.17	0.06	0.25	2.4	3271	OptionMetrics
Institutional ownership (in %)	50.7	27.5	0	56.8	100	3271	CDA/Spectrum 13F
Fixed capital (in $m)	1012.5	3913.1	0.04	105.2	84101	3271	Compustat
Employees (in 000s)	15.2	34.8	0.01	2.4	371.7	3271	Compustat
Sales (in $m)	3968.04	11841.2	0.004	494.4	171652	3271	Compustat
Firm age	22.5	15.5	3	16	55	3271	Compustat
R&D (in $m)	286.9	811.3	0	39.4	12183	3271	Compustat
1 - Lerner index	0.86	0.04	0.76	0.87	0.96	3271	Compustat
Profits/Assets	0.09	0.17	-1.37	0.12	0.62	3271	Compustat
CEO age	55.6	7.6	32	56	89	1996	ExecuComp
CEO vega (in $000s)	157.6	292.0	0	73.73	4578.0	1845	ExecuComp
CEO delta (in $000s)	761.9	1695.4	0	295.1	34647.1	1845	ExecuComp
Governance index	9.04	2.7	2	9	16	921	RiskMetrics and Gompers et al. (2003)

As a preamble to our main analysis, we provide results of non-parametric regressions that consider the relationship between our innovation measures and options trading. Figure 1.1 presents the results. In both panels, we show a line for the local linear regression estimated by the lowest smoother with a bandwidth of 0.8. Panel A displays the non-parametric relationship between the natural logarithm of (one plus) the number of patents

[14]Descriptive statistics for all other variables used throughout the course of our study are in the Appendix, Table A.16.

granted (unweighted patent counts) and the natural logarithm of options volume. Panel B replicates the graph but uses our primary measure of innovation, the natural logarithm of (one plus) forward citation-weighted patent counts. As can be seen, the correlation between innovation and options trading is clearly positive and appears to be monotonically increasing across options volume.

Panel A. Patent counts Panel B. Cite-weighted patent counts

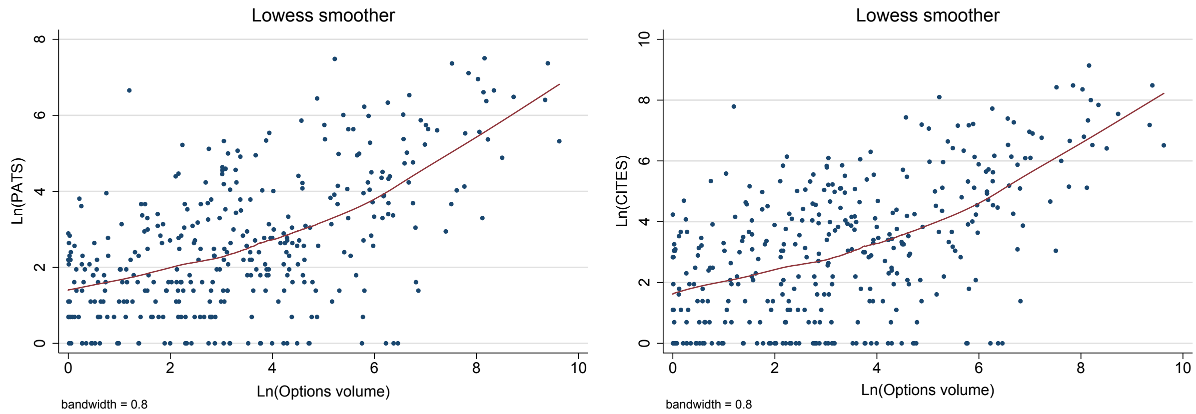


Figure 1.1: Non-parametric regressions of innovation and options volume

Notes: This figure presents non-parametric (local linear) regression of firms' unweighted patent counts (*PATS*) and annual options volume (Panel A) and firms' patents weighted by the number of forward citations (*CITES*) and annual options volume (Panel B). The graph is from 2000, the middle of our sample period.

1.3.3 Specification

Our main econometric models focus on the relationship between future cite-weighted measures of innovative activity and options trading. We estimate the following model using ordinary least squares (OLS):

$$Y_{i,t} = \alpha + \beta O_{i,t} + \gamma Z_{i,t} + \delta_t + \lambda_i + \varepsilon_{i,t} \tag{1.1}$$

where i indexes firms and t indexes time. The dependent variable, $Y_{i,t}$, is the natural logarithm of (one plus) the number of cite-weighted patents. The options trading measure, $O_{i,t}$, is measured for firm i over its fiscal year t as the logarithmic transformation of the options volume, although similar results are also obtained using the untransformed variable. Because both innovation and options activity are in logarithmic form, the coefficient on O gives us the elasticity of innovation to options trading. δ_t are time dummies that account for inter-temporal variation that may affect the relationship between options trading and innovation and λ_i is a firm fixed effect that controls for unobserved time-invariant firm heterogeneity. Because innovation metrics are likely to be autocorrelated over time, all of our models will allow the standard errors to have arbitrary heteroskedasticity and autocorrelation (i.e., clustering standard errors by firm). The vector $Z_{i,t}$ contains a range of control variables. Specifically, in our main regressions, we condition on firm size (*Sales*), capital-labour ratio (K/L) and deflated *R&D stock*, as suggested by the literature on patent production functions (e.g., Pakes and Griliches, 1980; Hausman, Hall, and Griliches, 1984). The model of Aghion, Van Reenen, and Zingales (2013) shows that innovative activities are affected by institutional ownership; we include the percentage of shares held by institutional investors (*InstOwn*). We also control for firm's *age* in the base model, measured as the number of years since the inclusion of the firm in Compustat.

When a firm's *R&D stock* is included in Z, we can interpret the equation as a "production function" that relates past R&D investments to innovative outputs. It follows that in this specification, β gives us the effect of options trading activity on the productivity of R&D, measured by forward cite-weighted patent counts per R&D dollar invested. Note that we also estimate models that omit the R&D stock from Z, and hence β indicates the combined impact of changes in R&D stocks and innovative productivity.

Finally, λ_i, the fixed effects, is introduced into the models using the "pre-sample mean scaling" estimator of Blundell, Griffith, and van Reenen (1999). Essentially, we exploit the fact that we have a long pre-sample history of a firm's innovative activities and construct pre-sample averages of the dependent variables.[15] This initial condition can proxy for unobserved heterogeneity if the first moments of the variables are stationary. Monte Carlo simulations show that this pre-sample mean scaling estimator performs well compared to alternative econometric estimators for dynamic panel data models with a long panel for innovations but only a short panel for the explanatory regressors.

1.4 Empirical Results

Table 1.2 presents our first set of regression results. Columns 1 through 4 report the OLS estimates with the dependent variable *Ln(1+CITES)*: the natural logarithm of (one plus) the number of citation-weighted patents for issued patents applied for in year t. Due to the count-based nature of citation and patent data, we also use count-based regression models, such as the Negative Binomial (NB). Columns 4 through 8 report NB regressions. Across all the columns of Table 1.2, the coefficient estimates on *Ln(Optvol)* are positive (ranging between 0.118 and 0.244) and both economically and statistically significant. For example, the coefficient of 0.118 in column 4 suggests that a 200% increase in the dollar volume of options traded (e.g., from the median of \$8.5 million to \$25.5 million) is

[15]We estimate from 1996 and use the information on patenting between 1976 and 1995 to construct the pre-sample means.

associated with a 24% increase in cite-weighted patents (e.g., from the median of 15 to 19).[16]

We begin in column 1 with OLS regressions of *Ln(1+CITES)* on options trading with controls for *InstOwn*, *Ln(K/L)*, *Ln(Sales)*, *Ln(Age)*, four-digit industry dummies and time dummies. Consistent with the bivariate relationships in Figure 1.1, there is a positive and significant association between innovation and options volume. Column 2 includes the controls for fixed effects (which are highly significant) and these substantially reduce the coefficient on *Ln(Optvol)* from 0.232 to 0.148. In columns 1 and 2, the options volume coefficient measures the combined impact of changes in R&D productivity (more innovative output per dollar of R&D invested) and innovative intensity (greater spending on innovation). In column 3, we add the natural logarithm of each firm's deflated R&D stock, and hence the equation becomes a production function, where β indicates the innovative premium of options trading per dollar of R&D. As expected, the coefficient on *Ln(R&D stock)* shows a very robust positive association with patent citations. The coefficient on options volume also declines by approximately 32%, from 0.232 to 0.158, indicating that the main effect of options trading operates by impacting R&D productivity rather than by stimulating more R&D spending. Column 4 presents the full model, which includes the controls for fixed effects. As before, this reduces the options volume coefficient from 0.158 to 0.118.[17] The final four columns of Table 1.2 repeat the main OLS specifications but use NB models. Our findings are similar.

[16] In the sample period between 1996 and 2004 the trading volume for our firms rose by 188%, so 200% is a reasonable change to consider.

[17] The results are similar if we replace the Blundell, Griffith, and van Reenen (1999) controls for fixed effects with the Hausman, Hall, and Griliches (1984) approach. For example, in an identical specification to our main model in column 4, the coefficient (standard error) on *Ln(Optvol)* is 0.091 (0.034).

Table 1.2: Innovation and options volume

Method	OLS				NB			
Dependent Var.	Ln(1+CITES)				CITES			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Optvol)	0.232***	0.148***	0.158***	0.118***	0.244***	0.157***	0.163***	0.121***
	(0.027)	(0.022)	(0.025)	(0.022)	(0.026)	(0.022)	(0.026)	(0.021)
InstOwn	-0.229	-0.139	-0.046	-0.048	-0.211	-0.127	0.093	0.064
	(0.200)	(0.165)	(0.178)	(0.158)	(0.228)	(0.188)	(0.206)	(0.177)
Ln(K/L)	0.065	-0.004	0.102*	0.026	0.224**	0.075	0.260***	0.110
	(0.070)	(0.056)	(0.061)	(0.053)	(0.097)	(0.074)	(0.083)	(0.067)
Ln(Sales)	0.395***	0.266***	0.141***	0.140***	0.399***	0.293***	0.159***	0.142***
	(0.046)	(0.036)	(0.053)	(0.041)	(0.044)	(0.035)	(0.048)	(0.040)
Ln(Age)	0.115	-0.037	-0.051	-0.109	-0.062	-0.136	-0.266***	-0.213**
	(0.106)	(0.087)	(0.096)	(0.083)	(0.109)	(0.098)	(0.101)	(0.094)
Ln(R&D stock)			0.462***	0.262***			0.486***	0.302***
			(0.058)	(0.046)			(0.047)	(0.042)
Firm fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3271	3271	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS and NB panel regressions of firms' patents weighted by the number of forward citations (*CITES*) on options volume (*Optvol*) and other firm-level control variables. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies and time dummies. The time period is 1996 – 2004 (with citations up to 2006); fixed effects are based on including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.4.1 Alternative innovation measures

In Table 1.3, we ask whether our results are driven by greater innovation output (more patents) or greater innovation input (more R&D expenditure), and whether our results are robust to the exclusion of self-citations. We find support for all three effects.

Columns 1 and 2 report the regression results from replacing the dependent variable of cite-weighted patents with raw patent counts. We observe a pattern for the coefficient of options trading activity that is very similar to that in our baseline models (i.e., columns 3 and 4 of Table 1.2). We observe a positive and significant coefficient estimate of *Ln(Optvol)* that falls substantially after we introduce time-invariant, firm-specific innovation determinants into the regressions. The effect, however, remains economically and statistically significant. For example, the coefficient estimate in column 2 implies that an increase in options volume by 200% leads to roughly 2 additional patents filed for by a median firm in our sample. Given that a median firm files for 7 patents, this is a significant increase.

The middle two columns examine the association between options trading activity and R&D investment. We remove the deflated *R&D stock* from this specification because we are interested in inputs and rely instead on a conditional fixed-effects estimator. In columns 3 and 4, we find that options volume has a significant and positive association with firm R&D investment, although the magnitude of this effect becomes smaller than that for cite-weighted patents after we add fixed effects. Thus, focusing on R&D as the only measure of firm innovativeness may underestimate the importance of options trading.

Columns 5 and 6 of Table 1.3 show that the coefficient estimates on *Ln(Optvol)* continue to be positive and significant at the 1% level when we remove self-citations and re-estimate Eq. (1.1) with the dependent variable replaced by the number of patents weighted by non-self citations. We find this last result important because the interpretation that our results are mainly driven by pure managerial signalling behaviour (as opposed to pushing the firm towards more innovation) is difficult to reconcile with our finding that firms with higher levels of options trading activity generate more forward citations in general and receive more forward citations from other firms in particular (e.g., as compared to an increase in patenting).

Table 1.3: Alternative innovation measures and options volume

Dependent Var.	Ln(1+PATS)		Ln(1+XRD)		Ln(1+NS_CITES)	
Method: OLS	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.176***	0.158***	0.268***	0.102***	0.198***	0.157***
	(0.029)	(0.024)	(0.029)	(0.014)	(0.032)	(0.028)
InstOwn	-0.273*	-0.213	-0.283*	0.017	-0.025	-0.045
	(0.154)	(0.140)	(0.162)	(0.102)	(0.167)	(0.149)
Ln(K/L)	0.101**	0.043	-0.060	0.080**	0.094*	0.032
	(0.051)	(0.044)	(0.062)	(0.031)	(0.056)	(0.052)
Ln(Sales)	0.144***	0.118***	0.516***	0.240***	0.133***	0.130***
	(0.046)	(0.033)	(0.041)	(0.034)	(0.049)	(0.039)
Ln(Age)	0.010	-0.028	0.087	0.605***	-0.022	-0.094
	(0.079)	(0.069)	(0.078)	(0.151)	(0.091)	(0.080)
Ln(R&D stock)	0.432***	0.205***			0.431***	0.255***
	(0.056)	(0.043)			(0.056)	(0.045)
SIC four-digit dummies	Yes	Yes	Yes	n/a	Yes	Yes
Firm fixed effects	No	BGV	No	Yes	No	BGV
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS regressions of firms' unweighted patent counts (*PATS*), R&D expenditure *(XRD)* and firms' patents weighted by the number of non-self forward citations *(NS_CITES)* on options volume (*Optvol*) and other firm-level control variables. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of time dummies. The time period is 1996 – 2004 (with non-self citations up to 2006); *BGV* fixed effects controls use the Blundell, Griffith, and van Reenen (1999) pre-sample mean scaling estimator. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.4.2 Robustness checks

We conduct a rich set of basic robustness tests for our baseline results and discuss the details of these tests in the Appendix. To summarize, we find that the positive effect of options trading activity on innovation continues to increase monotonically, is preserved in the two sub-periods during and after the Internet bubble, is robust to alternative proxies for R&D inputs, lagged options volume, and alternative econometric models that deal with the right-skewed and non-negative nature of patents data.

To provide additional insights, we conduct several tests related to our main prediction. To save space, these results are tabulated in the Appendix. First, we check if our results are robust to the inclusion of additional (financial) control variables. Although our approach is to condition on a wide range of firm characteristics (and fixed effects), one could object that this does not adequately control for observable omitted variables variables. For example, there may be concerns that our regressions omit the variable of firm's market value, which is correlated with the number of citations (Hall, Jaffe, and Trajtenberg, 2005). Given that fundamentals that increase firm's value may also increase innovation and given that informed traders may be more likely to trade firms with higher growth opportunities, this may produce a spurious upward bias on the coefficient on options volume.[18] Similarly, He and Tian (2013) and Fang, Tian, and Tice (2014) show that analyst coverage and stock liquidity are important determinants of firm innovation. To address such concerns, in the Appendix, Table A.8, we augment our main specification by including stock illiquidity (*Illiquidity*), leverage (*Leverage*), stock market-based firm value (*Tobin's Q*), profitability (*ROA*), capital expenditures (*Capex*) and analyst coverage (*Analyst coverage*). However, the coefficients on *Ln(Optvol)* continue to be positive and significant at the 1% level, and the magnitude of the coefficient declines only slightly from the baseline model (i.e., from 0.118 to 0.110).[19]

Another concern might be that our results are affected by firms' external knowledge sourcing behaviour. Because of the increased complexity of the technological and scientific developments in our focal industries, firms cannot rely solely on internal R&D; they need to (and do) source knowledge externally in order to enhance the performance of their innovation process (Cassiman and Veugelers, 2006). Simultaneously, when prices are more efficient, managers can extract more information from the market so that they can better assess the quality and potential outcomes of external knowledge acquisition activities.[20] To account for this, we include three alternative variables in our set of controls. Specifically, we use the *frequency* to which firms engage in R&D collaborations (i.e., the number of alliances and joint ventures reported in SDC), the *intensity* to which firms have sourced external knowledge (i.e., the number of jointly owned patents divided by the total number of patents) and the *acquisition* of innovative target firms (i.e., acquisition expen-

[18]The correlation between *Tobin's Q* and *Ln(Optvol)* is positive (0.262) and significant at the 1% level.

[19]In unreported results, we also consider return volatility, measured by the annualized standard deviation of daily returns. However, and consistent with the notion in Roll, Schwartz, and Subrahmanyam (2009), the return volatility variable is not significant. For example, in an identical specification to column 1 Table A.8 in the Appendix, the coefficient (standard error) on return volatility is 0.017 (0.012), while the coefficient on *Ln(Optvol)* remains positive and significant at the 1% level, with the magnitude of the estimate almost identical to the one reported above.

[20]Prior studies provide evidence consistent with this argument. For example, Luo (2005) finds that the positive correlation between announcement date return and the completion of mergers can be attributed to insiders' learning from outsiders after controlling for common information.

diture normalized by total assets). We report the results in the Appendix, Table A.9. We find that the coefficients on *Ln(Optvol)* continue to be of a very similar magnitude as in column 4 of Table 1.2 (except when controlling for collaboration frequency because we only consider firms that have some information in SDC) and continue to be significant at the 1% level. Overall, this provides reassurance that our findings are primarily related to internal R&D investment decisions.

Third, we argue that information asymmetries between the firm and market participants are especially challenging in R&D-intensive industries (which is one reason why this is our sample of interest). This is because the nature of firms' core activities is knowledge-based and highly opaque, and the fact that there could be a substantial cost to revealing information to their competitors reduces the quality of the signal they can make about their innovative activities (Bhattacharya and Ritter, 1983; Anton and Yao, 2002). Thus, if what we are capturing is related to the informational benefit from options to a firm in reducing asymmetric information problems related to R&D, then this should matter more for firms that are active in R&D-intensive industries, relative to cases where such problems are less (or not at all) present. To show this, we begin by identifying firms with positive options volume and non-missing data on all other variables that operate in non-R&D-intensive industries, defined as those that are located in the OECD classification (based on R&D intensities) of low-tech industries (OECD, 2011).[21] Given that these firms are very different from our focal firms, we then apply a matching procedure that relies on a nearest neighbour matching of propensity scores (estimated as a function of all firm characteristics, including fixed effects). After restricting the sample to common support, we are left with a panel of 1,453 firm-years in both groups. In Column 1 of Table A.10 in the Appendix, we estimate our main specification on the matched sample, adding a dummy variable for R&D-intensive firms (= 1;0 = non-R&D-intensive). The coefficient on *Ln(Optvol)* remains positive (0.120) and statistically significant at the 1% level, while the coefficient on the dummy is also positive (1.219) and significant at the 5% level. In column 2, we add the interaction of this dummy variable with options volume. The estimates show that the interaction term, *Ln(Optvol) x Dummy for high-tech*, is 0.186 and highly significant, as expected. Most interesting, however, the coefficient on *Ln(Optvol)* goes towards zero (0.004) and becomes insignificant once the interaction term is included. Taken literally, this indicates that there is no effect of options trading activity on innovation in non-R&D-intensive industries, which is broadly consistent with the story we present. For robustness, we also split the sample. In column 3, in R&D-intensive industries, the coefficient on options volume is large, positive and significant at the 1%

[21] According to industrial codes of ISIC Rev.3 (NACE Rev.1 in Europe), OECD (2011) classifies manufacturing industries in four subgroups of high-technology, medium high-technology, medium low-technology and low-technology based on the technology intensity and level of R&D used in these industries.

level, whereas in column 4, in non-R&D-intensive industries, the coefficient is smaller and insignificant (0.144 versus 0.070).[22]

Finally, we perform a small event study that examines the effect of initial option listings on firms' innovation performance. To do so, we focus on the sub-sample of firms who appear for at least two years before and after the listing event. After excluding firms with multiple listings, we are left with a set of 93 events during the period between 1998 and 2002. Next, we proceed to construct a dummy variable, *Post*, that equals one for the post-event period and zero for the pre-event period. In column 1 of Table A.11 in the Appendix, we augment Eq. (1.1) by including *Post*. The within-firm estimator is of 0.370 and significant at the 5% level. In terms of economic significance, this suggests that option listing is associated with a 37% increase in innovation in subsequent periods. In our second diagnostic test, we examine the dynamics of innovation in the years around the listing event. We use a window of eight years and include in our main specification a set of dummy variables for the three years prior to the year when the firm was listed and four years after the firm was listed (year zero is the omitted category). Figure 1.2 presents the results. Panel A depicts the within-firm changes in the raw number of patents and panel B shows the changes in cite-weighted patents. In both figures, we find that there is little effect in the first year after listing but in the following years innovation increases substantially by about 64% (measured by cite-weighted patents) with respect to the listing year.

Two conclusions emerge from this analysis. First, it mitigates concerns that options trading is endogenous due to reverse causality. This is a particular concern when studying innovation output because it is difficult to be addressed by lagging the explanatory variable. Even if innovation is regressed on lagged options volume (as done in the Appendix, Table A.3), it may be that lagged innovation causes lagged options volume and also causes current innovation. Because we only consider the first listing (based on the data available to us), it cannot be caused by past listings. Second, the phenomenon of such delayed effects is consistent with the starting point of our theory, i.e., that the benefit from options is related to whether the market for the listed option has sufficient volume, because trading volume needs time to build.

[22]Clearly, as detailed in Section 1.3, this approach has the problem that patent and citation data are a less reliable indicator of innovation in low-tech industries. To address this, we experimented with different sub-samples of non-R&D-intensive firms such as including low- and/or medium low-tech industries and focusing only on those firms that have non-zero citations in more than 25% or 50% of the years they appear in our sample. Our finding, however, remains unaltered in all these tests. For example, re-doing the analysis using a matched sample of firms in low-tech industries which receive citations in more than 50% of the years yields the following results: we estimate a coefficient on the interaction term of 0.240 (standard error = 0.061) and a coefficient on options volume of -0.054 (standard error = 0.052) on this sub-sample of 2,192 observations. If we split this sub-sample into R&D- and non-R&D-intensive industries, the coefficient on *Ln(Optvol)* is large and significant only for firms in R&D-intensive industries (i.e., a coefficient of 0.130 with a standard error of 0.053 versus a coefficient 0.023 with a standard error of 0.056).

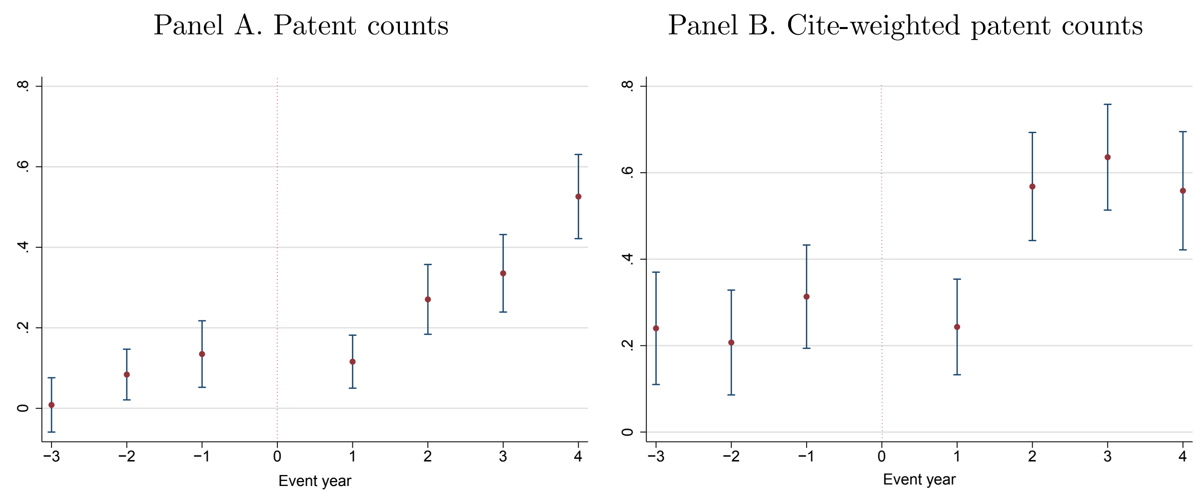


Figure 1.2: Innovation around the option listing event

Notes: This figure presents within-firm changes in (Panel A) firms' unweighted patent counts (*PATS*) and (Panel B) firms' patents weighted by the number of forward citations (*CITES*) in the years around the option listing event. The estimates are taken from the regressions reported in column 2 and 4 of Table A.11 in the Appendix.

1.4.3 Innovative direction

We make now use of alternative outcome variables to explore the idea that our baseline results are not simply driven by greater R&D productivity (more citations per R&D dollar) but are rather associated with different resource allocation decisions (or innovative directions). To perform such a test, we use three different measures of the direction of firms' innovative efforts: (i) an originality index on knowledge inputs, (ii) a measure of risk-taking behaviour and (iii) a proxy for innovative diversity. The results are shown in Table 1.4. As before, all models are estimated via fixed effects OLS panel regression using the pre-sample mean of the dependent variables. We lose some observations in these specifications because in the fixed effects estimator, we require a firm to have at least some information on the dependent variables in the 1996 – 2004 and pre-sample period.

In column 1, the outcome variable is an originality-weighted patent count. Originality, as defined in Hall, Jaffe, and Trajtenberg (2001), is essentially one minus a Herfindahl index of the concentration of backward patent citations across two-digit technological classes. We find that the coefficient estimate on *Ln(Optvol)* is positive and significant at the 1% level, suggesting that firms with more options trading activity make use of a more diverse set of knowledge.[23] In the next column, we use a measure of risk-taking behaviour, i.e., the standard deviation of forward citations received across patents. The

[23]The coefficient on options volume continues to be positive and significant when we use NB (coefficient of 0.157 with a standard error of 0.024) and Poisson (coefficient of 0.114 with a standard error of 0.050) specifications.

results show that there is a positive and meaningful relationship between *Ln(Optvol)* and *Ln(1+SD_CITES)*.[24] In column 3, we introduce the diversity of innovation activities, defined as one minus the Herfindahl index of the number of patents across classes. The patent portfolio includes all patents of a given firm over a three-year period. To control for the fact that some firms are engaged in so little innovation that it may not be meaningful to speak of a diverse (or concentrated) technological direction, we define a minimum threshold of five patents to filter out such low-innovation firms.[25] As shown, the estimate on *Ln(Optvol)* is positive and significant at the 10% level, suggesting that firms with more options trading activity file for a more diverse set of patents.[26]

Taken together, the results in Table 1.4 reveal that options trading appears to evoke a change in the direction of innovative efforts as opposed to merely increase the amount of R&D or patenting; on average, firms with more trading activity produce a more diverse and original set of activities and are characterized by an increasing willingness to take risk in their innovation process. We find these results especially intriguing because, intuitively, one may conjecture that the disciplinary feature of financial markets is to force managers to refrain from engaging in overly risky projects and to abandon creativity and diversity for efficiency purposes. We take these results as a starting point to explore the question of why this is the case in Section 1.5.

[24]Our findings are similar if we use the untransformed variable: the coefficient (standard error) on *Ln(Optvol)* is 0.319 (0.085) in column 2 of Table 1.4 when *Ln(1+SD_CITES)* is replaced by *SD_CITES*.

[25]We also consider alternate cut-off points of 10, 20 and 50 patents and obtain similar results.

[26]For robustness purposes, we apply an alternative modelling approach that accounts for the bounded nature of the dependent variable. Specifically, we employ a double-truncated Tobit model and find similar results for the coefficient on options trading (i.e., the coefficient is 0.008 with a standard error of 0.004).

Table 1.4: Innovation direction and options volume

Dependent Var.	Ln(1+ORIG.)	Ln(1+SD_CITES)	DIVERSITY
Method: OLS	(1)	(2)	(3)
Ln(Optvol)	0.154***	0.039***	0.007*
	(0.022)	(0.010)	(0.004)
InstOwn	-0.140	0.104**	0.028
	(0.127)	(0.052)	(0.023)
Ln(K/L)	0.032	-0.022	0.005
	(0.037)	(0.024)	(0.010)
Ln(Sales)	0.086***	0.010	0.027***
	(0.030)	(0.012)	(0.007)
Ln(Age)	-0.026	-0.082***	0.018
	(0.060)	(0.029)	(0.013)
Ln(R&D stock)	0.159***	0.041***	-0.003
	(0.040)	(0.012)	(0.006)
Observations	3245	2382	1526

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by originality (*ORIG.*), the standard deviation of forward citations across firms' patents (*SD_CITES*), and innovative diversity *(DIVERSITY)* on options volume (*Optvol*) and other firm-level control variables. See Hall, Jaffe, and Trajtenberg (2001) for a definition of originality. Diversity is defined as one minus the Herfindahl index of the number of patents across two-digit technological classes. Firms in columns: 542 in column 1, 455 in column 2, and 362 in column 3. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.4.4 Endogeneity of options trading

In this subsection, we address concerns that informed traders select firms on the basis of characteristics that are observable to them but not observable for us. For example, informed investors might decide to trade options on stocks when they anticipate an increase in innovation. Another problem might be that our measure of options trading activity is noisy. This is because intra-day execution prices are not available over the large sample period. Although this is mitigated by annual averaging, we are likely to underestimate the effect of options.[27]

We address these issues in a number of ways. First, we use matching estimators

[27] Another issue might be that different types of options provide different signals. Although we could bring additional data to examine the breakdown of call and put options with different times to maturity, there are no clear hypotheses. As Roll, Schwartz, and Subrahmanyam (2009) note, while it may be the case that managers are more likely to act on "good news" as opposed to "bad news", calls and puts can be bought and sold freely. Thus, in the absence of information on the signed order imbalance (data we unfortunately don't have), disaggregating calls and puts cannot be unambiguously linked to managerial investment decisions.

to calculate the average effect of having high levels of options volume on innovation. Second, we consider moneyness and open interest as two plausible exogenous instrumental variables and perform 2SLS estimations of the regressions in Table 1.2. Third, we test in Section 1.5 whether our results are consistent with additional predictions and empirical findings concerning the environment in which options trading should have differential effect on innovation.[28]

Propensity score matching

We start with propensity score matching (PSM) in order to see whether firms with high trading activity would have innovated at a lower rate had they not high trading activity. In the application that follows, we define a high (low) trading activity firm as a firm with options volume above (below) the yearly median in a given three-digit SIC industry. The PSM technique is based on the likelihood that an observation would be a high trading activity firm conditional on observables (Rosenbaum and Rubin, 1983, 1984). We use a probit specification to estimate the probabilities of being a high trading activity firm (= 1;0 = otherwise) on a comprehensive list of observable characteristics, including all the independent variables (including the additional controls), as well as fixed effects. We then use the predicted probabilities, or propensity scores (stratified by industry and year), from this probit estimation and perform the matching. As our main matching procedure we use nearest neighbour matching that allows each treated firm to be matched with multiple controls (i.e., 4, although our result are robust to any number of matches between 1 and 5), running the procedure with replacement. However, to ensure that the results are not sensitive to our choice of matching estimator, we also provide evidence from kernel and radius matching.

Table 1.5 reports the average treatment effect estimates. The average selection bias (not tabulated) across all specifications ranges from 3.4% to 4.7%, which means that the results are reliable. As can be seen, our findings are in line with those obtained in the previous panel regressions. For example, the results in columns 1 and 2 suggest that firms with high trading activity produce 70% more patents that subsequently generate 72% more citations per dollar of R&D than firms with low options volume, all significant at the 1% level. Overall, this suggest that the non-random assignment of high levels of options trading to more innovative firms (at least based on observables) does not explain our findings.

[28] An interesting context for the purpose of our study would be the use of a difference-in-differences estimator that relies on the exogenous variation in options trading generated by short selling bans or constraints imposed by regulators during the 2007 – 2009 crisis. Unfortunately we lack data on patenting during and after the crisis period.

Table 1.5: Propensity score matching

	Panel A: Nearest neighbour matching		Panel B: Kernel Matching		Panel C: Radius Matching	
	Ln(1+CITES)	Ln(1+PATS)	Ln(1+CITES)	Ln(1+PATS)	Ln(1+CITES)	Ln(1+PATS)
	(1)	(2)	(3)	(4)	(5)	(6)
High Optvol	0.722***	0.704***	0.468***	0.382***	1.155***	1.096***
vs.	(0.104)	(0.092)	(0.125)	(0.094)	(0.072)	(0.063)
Low Optvol						
Observations	2130	2130	2130	2130	2130	2130

Notes: This table presents estimates of differences in firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) between the treatment group (high levels of options volume) and the control group (low levels of options volume). The matched sample is constructed using a nearest-neighbour (Panel A), kernel (Panel B) and radius (Panel C) score matching with scores given by a probit model in which the dependent variable is a dummy variable that equals one if a firm has an options volume above the yearly median in a given three-digit industry. The propensity score is estimated using the following firm characteristics: *InstOwn*, *Ln(K/L)*, *Ln(Sales)*, *Ln(Age)*, *Ln(R&D stock)*, *Illiquidity*, *Leverage*, *Tobin's Q*, *ROA*, *Capex*, *Ln(Analyst coverage)* and fixed effects. Firms in columns: 525. Standard errors are obtained using 200 bootstrap replications (in parentheses). The time period is 1996 – 2004 (with citations up to 2006). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Instrumental variable

Our second approach to correct for the potential bias due to selection is an instrumental variable strategy. Roll, Schwartz, and Subrahmanyam (2009) propose two instrumental variables that are reasonably exogenous to the relationship between options volume and innovation: (i) moneyness (i.e., the average absolute difference between the stock's market price and the option's strike price) and (ii) open interest in the stock's listed options. We focus our analysis on moneyness while Table A.12 in the Appendix shows that our results are similar if we use the total open interest in the stock's listed options as alternative.

A good instrument is a variable that is correlated with options trading (this assumption can be tested) but uncorrelated with our dependent variables except through other independent variables. That is, the instrument should be a variable that can be excluded from the original list of controls without affecting the results. As Roll, Schwartz, and Subrahmanyam (2009) argue, there are several reasons that moneyness is related to options trading. First, informed traders may be more attracted to out-of-the-money (OTM) options because they offer the greatest leverage but uninformed agents may prefer in-the-money (ITM) options to avoid overly risky positions (Pan and Poteshman, 2006). Moreover, volatility traders would avoid deep ITM or OTM options, as the vega of such options is close to zero. Specifically, Chakravarty, Gulen, and Mayhew (2004) show that the trading volume by agents speculating on volatility tends to be concentrated in at-the-money (ATM) options. In sum, these arguments suggest that moneyness is related to options trading, although they do not establish an unambiguous direction. There is no reason to believe, however, that (unsigned) moneyness is linked to innovation in any intrinsic way because exchanges periodically list new options with strike prices close the stock's market price.

Given that there is no strong rationale for a mechanical link between moneyness and innovation, we use the average absolute moneyness as an instrument. This variable is measured as the yearly average of the daily absolute deviation of the exercise price of each traded option from the closing price of the underlying asset.[29] The correlation of this variable with options volume is 0.326, which suggests that the instrument is indeed related to options trading, and is consistent with that reported in Roll, Schwartz, and Subrahmanyam (2009). We implement the instrumental variable estimator using 2SLS.

Column 1 of Table 1.6 presents the first stage which regresses options volume on moneyness (and all other controls). As indicated by the simple correlation, we find that the instrument is positive and highly significant. Moreover, the first-stage F-stat for the "weak instrument rule of thumb" is strongly significant (and well above ten), which suggests that the hypothesis that the instrument can be excluded from the first-stage regressions is rejected and that the instrument is not weak. Columns 2 and 3 present the coefficient estimates for the second stage, where we control for endogeneity. Column 2 presents the results with cite-weighted patents as dependent variables. Consistent with the findings from the OLS specification, the coefficient estimate on *Ln(Optvol)* is positive and significant at the 1% level. In column 3, we find a very similar pattern using unweighted patents as the dependent variable. To provide additional support for the validity of our instrumental variable approach, we replicate the estimations of Table 1.6 using open interest as an additional instrument and rely on the Hansen J-statistic. Our instrument performs adequately in our tests (p-value = 0.93 and 0.31 in identical specifications of columns 2 and 3, respectively), indicating that we cannot reject the null hypothesis of instrument suitability.

To gauge the direction and magnitude of the bias due to the endogeneity of options trading, we can compare the OLS results from Table 1.2 with those obtained from the 2SLS regressions. Interestingly, the 2SLS coefficient estimates on *Ln(Optvol)* are considerably larger (i.e., more positive) than those of the OLS estimates, although the estimates from both approaches are in the same direction and statistically significant.[30] This OLS bias towards zero could be because options trading is measured with some error or because omitted variables simultaneously make firms innovative and more attractive to informed

[29]For traded option k on stock j for day d, the absolute deviation is $|\text{Ln}(\text{Price}_{j,d}/\text{Strike}_k)|$. This is averaged over all k and d within a year t for each stock j. Options without trades are not included in the calculation of the moneyness variable. We obtain very similar results if we use volume-weighted average annual moneyness for each stock j, where each option's moneyness is weighted by the proportion of total option volume for stock j contributed by that option.

[30]For robustness, we also consider the instrumental variable estimator by using the control function approach (Blundell and Powell, 2004). The coefficient estimate (standard error) on *Ln(Optvol)* in the control function estimation was also above the ordinary Poisson estimate (see Appendix, columns 4 and 8 of Table A.1): 0.188 (0.087) for *CITES*, and 0.165 (0.065) for *PATS*.

traders. The attitudes and beliefs of CEOs could be an example of such omitted variables. For instance, overconfident CEOs could attract more informed traders, while simultaneously, they could also be more likely to pursue innovation that results in more patents and citations (Hirshleifer, Low, and Teoh, 2012)

Table 1.6: Innovation and options volume – Moneyness as instrumental variable

Method	OLS (first stage)	2SLS (second stage)	
Dependent Var.	Ln(Optvol)	Ln(1+CITES)	Ln(1+PATS)
	(1)	(2)	(3)
Ln(Optvol) *(instr.)*		0.180***	0.162***
		(0.057)	(0.062)
InstOwn	1.051***	-0.033	-0.259*
	(0.224)	(0.179)	(0.134)
Ln(K/L)	-0.365***	0.095	0.061
	(0.080)	(0.067)	(0.050)
Ln(Sales)	0.595***	0.151**	0.103**
	(0.052)	(0.063)	(0.042)
Ln(Age)	-0.322***	-0.059	-0.014
	(0.123)	(0.101)	(0.071)
Ln(R&D stock)	0.361***	0.471***	0.194***
	(0.054)	(0.071)	(0.050)
Ln(Moneyness)	1.343***		
	(0.155)		
Observations	3271	3271	3271

Notes: This table presents estimates of 2SLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*) and other firm-level control variables, with the average absolute moneyness *Ln(Moneyness)* as instrumental variable. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.5 Possible Mechanisms

Our evidence thus far is consistent with the implication of our leading hypothesis that options trading enhances firm innovation, even after accounting for potential endogeneity concerns. In this section, we turn to the last part of our analysis and discuss potential underlying mechanisms through which this may occur. It is of course challenging to provide definite proof, and hence our tests are only suggestive.

On a broad view, we suggest two possible explanations for our results. The first is that managers prefer the "quiet" life as proposed by Hart (1983), Schmidt (1997) and

Bertrand and Mullainathan (2003) and that the increased price informativeness induced by options trading serves as a monitoring mechanism that forces managers to invest in innovation if they are a priori reluctant to do so. Alternatively, the positive association between options trading and innovation can also be attributable to career concerns. Most prominently, Aghion, Van Reenen, and Zingales (2013) recently extend the Holmström (1989, 1999) career concern model in the context of institutional investors (i.e., blockholders) and innovation. Based on the observation that managers concerned with their reputations in the labour market have incentives to take actions that boost current earnings and the firm's current stock price (Narayanan, 1985); the authors' findings suggest that the presence of institutional investors "protects" managers against the reputational risk associated with long-term investments in innovation. Because of their informational advantage, they have the ability to assess managerial efforts in innovation independent of potential bad profit realizations in the short run. This, in turn, provides incentives for the manager to forgo short-term profits and to invest in innovation. To the extent that the previous literature, both theoretical and empirical, argues that options increase the amount of private information conveyed by prices (e.g., Cao, 1999; Chakravarty, Gulen, and Mayhew, 2004; Pan and Poteshman, 2006; Roll, Schwartz, and Subrahmanyam, 2009; Hu, 2014), we may expect that this rationale also applies in the context of active options markets.

To understand the extent to which the aforementioned stories might explain our findings, we implement several tests concerning the environments in which options trading activity should have differential effects on innovation. First, we examine whether the effect of options trading on innovation depends on product market competition. The quiet life story suggests that the effect of options trading on innovation is weaker in highly competitive environments because stronger competition increases the threat of bankruptcy, which induces the manager to work harder to avoid liquidation and to keep his job (Hart, 1983; Schmidt, 1997). In contrast, if informed agents serve as a "shield" that protects managers, this effect should be more pronounced when the degree of product market competition is higher. This is because competition reduces the chances of success and hence increases the reputational risk faced by managers if they do so. Second, we investigate how innovation varies with options volume and managerial entrenchment. As Ferreira, Ferreira, and Raposo (2011) show, a disciplining takeover is more likely when prices are more efficient. Thus, an implication of our preceding discussion is that if managers prefer the quiet life, the beneficial effect of options should be stronger when managers are more "entrenched". Specifically, if a firm adopts a large number of takeover defenses, it might become partially insulated from the market for corporate control. In such cases, the takeover market cannot play an effective disciplinary role, and managers have greater ability to shirk. Moreover, if shareholder rights are restricted (i.e., the manager has more bargaining power against shareholders), the CEO will also be more entrenched. Third, if career concerns are the driving force behind this relationship, the effect of options trading on innovation should be stronger for younger CEOs because they are more concerned with their careers, and to boost their careers, they are likely to engage in myopic behaviour. Gibbons and Murphy (1992) show that implicit incentives from career concerns are much more substantial for younger managers. Holmström (1999) notes that when managerial ability is initially unknown and managerial effort is unobservable, young managers will overwork to benefit their future careers. Thus, there should be little managerial slack for younger CEOs. As before, under the quiet life story, options trading should have less of an effect when managers are younger, while under the career concerns story, the impact of options trading on innovation should be stronger when managers are younger.

1.5.1 Product market competition

Table 1.7 presents several results related to the interaction between options volume and product market competition. To measure product market competition, we use the inverse Lerner index (as in Aghion, Bloom, Blundell, Griffith, and Howitt (2005)), defined as one minus the median gross margin across all firms in the entire Compustat database with the same three-digit industry SIC as the focal firm. Our main model allows this measure to vary over time but we also consider its time-invariant form.

The first column reproduces our baseline results (column 4 of Table 1.2) and introduces the time-varying measure of product market competition. In this specification, the coefficient estimate on competition is positive and statistically significant (more competition yields more innovation), while the coefficient on *Ln(Optvol)* remains positive and significant.[31] Column 2 includes the interaction term between options trading and product market competition which is positive and significant at the 1% level, as predicted by the career concern hypothesis. In columns 3 and 4, we then replace the dependent variable with raw patent counts and repeat the analysis. We observe similar patterns for the interaction term. For robustness, columns 5 and 6 repeat the same specifications as above but restrict the inverse Lerner index to be constant over time. This yields similar results. Note that we are not able to estimate the main effect of competition in this model because the measure is collinear with industry effects.

[31] In line with Aghion, Bloom, Blundell, Griffith, and Howitt (2005), we find some evidence of an inverted U-shaped relationship between innovation and product market competition. If we include a term in the square of the inverse Lerner index, it is negative, whereas the linear term remains positive. This quadratic term is insignificant, however, with a coefficient estimate of -32.205 and a standard error of 28.076.

Table 1.7: Innovation and options volume – Product market competition

Measure of competition	Varies over time				Constant over time	
Dependent Var.	Ln(1+CITES)		Ln(1+PATS)		Ln(1+CITES)	Ln(1+PATS)
Method: OLS	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)		2.467***		1.962***	3.334***	2.579***
x Competition		(0.469)		(0.249)	(0.284)	(0.319)
Ln(Optvol)	0.169**	0.169***	0.160**	0.160***	0.165***	0.156***
	(0.046)	(0.026)	(0.037)	(0.020)	(0.029)	(0.022)
Competition	5.834***	5.054**	6.852***	6.229***		
(1 − Lerner)	(1.248)	(2.024)	(0.649)	(1.213)		
InstOwn	-0.035	0.017	-0.212	-0.171	0.027	-0.165
	(0.229)	(0.183)	(0.224)	(0.181)	(0.182)	(0.181)
Ln(K/L)	0.026	0.041	0.042	0.053	0.047	0.058
	(0.100)	(0.080)	(0.088)	(0.071)	(0.081)	(0.071)
Ln(Sales)	0.132	0.147*	0.121	0.133**	0.151*	0.135**
	(0.093)	(0.085)	(0.069)	(0.063)	(0.086)	(0.063)
Ln(Age)	-0.107	-0.089	-0.032	-0.017	-0.077	-0.007
	(0.148)	(0.111)	(0.092)	(0.069)	(0.112)	(0.071)
Ln(R&D stock)	0.251***	0.243***	0.202***	0.195***	0.244***	0.197***
	(0.048)	(0.052)	(0.035)	(0.042)	(0.052)	(0.041)
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*), product market competition (*Competition*), their interaction and other firm-level control variables. Firms in columns: 548. Robust standard errors are clustered at the three-digit industry level (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006). Product market competition is constructed as 1 − Lerner index where Lerner is calculated as the median gross margin from the entire Compustat database in the firm's three-digit industry. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.5.2 Managerial entrenchment

Table 1.8 analyses the interaction between options trading and managerial entrenchment. To measure the degree of managerial entrenchment, we use the "Governance Index" (G-Index) introduced by Gompers, Ishii, and Metrick (2003). It consists of twenty-four corporate governance provisions and is based on firm-level corporate governance provisions and firms' governing state corporate law statutes. We obtain this information from RiskMetrics. Because this covers S&P 1500 firms in 1998, 2000, 2002 and 2004, our sample size declines in this analysis. A higher G-Index score indicates more restrictions on shareholder rights or a greater number of anti-takeover measures.

Our evidence is consistent with the findings in Table 1.7. In line with the career concern hypothesis (and in contrast to the quiet life hypothesis), the positive association between options trading and innovation is stronger when managers are less entrenched. Specifically, the interaction between options volume and managerial entrenchment in column 2 of Table 1.8 generates a significantly negative coefficient estimate of -0.028 (significant at 5%) while the main effect of *Ln(Optvol)* remains positive and statistically significant at the 1% level.

For robustness purposes, Table A.13 in the Appendix investigates the interaction between options trading and the "Entrenchment Index" (E-Index) (Bebchuk, Cohen, and Ferrell, 2009). The findings are similar.

Table 1.8: Innovation and options volume – Managerial entrenchment

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Ln(Optvol)		-0.028**		-0.018
x G-Index		(0.014)		(0.013)
Ln(Optvol)	0.179***	0.170***	0.159***	0.154***
	(0.032)	(0.032)	(0.029)	(0.029)
G-Index	0.019	0.045	0.010	0.026
(Governance Index)	(0.033)	(0.033)	(0.029)	(0.028)
InstOwn	-0.067	-0.094	-0.031	-0.048
	(0.191)	(0.191)	(0.169)	(0.168)
Ln(K/L)	0.152**	0.151**	0.109*	0.109*
	(0.077)	(0.076)	(0.063)	(0.063)
Ln(Sales)	0.111**	0.109**	0.123***	0.122***
	(0.050)	(0.049)	(0.046)	(0.046)
Ln(Age)	-0.184**	-0.169*	-0.114	-0.105
	(0.092)	(0.092)	(0.085)	(0.085)
Ln(R&D stock)	0.111***	0.115***	0.103***	0.106***
	(0.031)	(0.030)	(0.027)	(0.027)
Observations	921	921	921	921

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*), managerial entrenchment (*G-Index*), their interaction and other firm-level control variables. Firms in columns: 331. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of three-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The G-Index is an average of 24 provisions in the firm's charter (see Gompers, Ishii, and Metrick, 2003). The measure is based on data from RiskMetrics in 1998, 2000, 2002 and 2004. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.5.3 CEO age

Finally, we examine how CEO's age alters the effect of options trading on innovation. To capture CEO's age, we extract information from ExecuComp. As before, this database covers firms in the S&P 1500 so we are left with a sub-sample. Under the career concern hypothesis, we expect that the effect of options trading on innovation should be more pronounced for younger CEOs. If anything, it should increase the impact because career concerns are stronger when managers are further from retirement, as that increases the returns from influencing the market's belief about their abilities. For this reason, investment decisions of younger CEOs should to be more affected by their career concerns than those of older CEOs.

Table 1.9 presents evidence on the interaction between CEO's age and options trading on innovation. The estimate in column 2 of Table 1.9 confirms our conjecture and shows a negative coefficient on the interaction term (-0.368) that is statistically significant at the 5% level. Consistent with our earlier findings, the coefficient on *Ln(Optvol)* continues to be positive and significant at the 1% level, while the coefficient on *CEO age* is negative (older CEOs are less innovative), although the effect is not significant.

Table 1.9: Innovation and options volume – CEO age

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Ln(Optvol)		-0.368**		-0.163
x Ln(CEO age)		(0.162)		(0.147)
Ln(Optvol)	0.166***	0.162***	0.161***	0.159***
	(0.044)	(0.043)	(0.035)	(0.035)
Ln(CEO age)	-0.245	-0.113	-0.309	-0.250
	(0.389)	(0.375)	(0.339)	(0.323)
InstOwn	0.005	0.036	-0.026	-0.012
	(0.227)	(0.226)	(0.202)	(0.203)
Ln(K/L)	0.050	0.045	0.097	0.095
	(0.083)	(0.082)	(0.068)	(0.068)
Ln(Sales)	0.299***	0.301***	0.241***	0.241***
	(0.068)	(0.069)	(0.060)	(0.060)
Ln(Age)	-0.155	-0.149	-0.086	-0.084
	(0.132)	(0.132)	(0.113)	(0.113)
Ln(R&D stock)	0.151***	0.155***	0.126**	0.128**
	(0.051)	(0.052)	(0.051)	(0.052)
Observations	1996	1996	1996	1996

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*), *CEO age*, their interaction and other firm-level control variables. Firms in columns: 337. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). CEO age is based on data from ExecuComp over the period 1996 – 2004. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.5.4 Profitability

To strengthen our explanation, in this section, we analyse one additional economic mechanism that is supposed to directly create greater market pressure imposed by investors – a decline in profitability. Specifically, Kothari (2001) finds that financial reporting conveys substantial information to outsiders about firm performance that significantly influences market expectations and stock prices. Moreover, survey evidence reveals that profitability is the most important externally reported performance measure and that the majority of managers are willing to cut discretionary spending (e.g., R&D) to meet or exceed benchmarks (Graham, Harvey, and Rajgopal, 2005). Thus, the short-term pressure imposed by external agents might be substantially more pronounced for firms with earnings that reflect decreasing profitability because investors are more likely to exit based on this negative information and the stock price may drop. Managers in these firms are also at more risk of being fired because boards aggressively fire CEOs for lower performance (Jenter and Lewellen, 2014). In sum, if it is true that options trading activity shields managers from short-term market pressures (and the risk of being fired), we expect that the positive effect of options trading is magnified for firms with a decline in profitability.

Table 1.10 shows the results. In column 1, we regress the citations-weighted patent count on the lagged change in profitability (adjusted by assets) and options trading (and all the other controls). We find that higher profitability growth has a negative association with innovation, but the effect is not significant, while the coefficient estimate on *Ln(Optvol)* remains positive and significant. Column 2 interacts the profitability variable with options volume. The coefficient on this interaction is negative and significant at the 1% level, suggesting that innovation is more sensitive to options trading when firms' profitability growth is lower. As before, columns 3 and 4 of Table 1.10 present the robustness test by replacing the dependent variable with simple patent counts. Although we observe a similar pattern, the coefficient on the interaction term is insignificant. This is interesting, however, because at face value, this result combined with the insignificant interactions in Table 1.8 and 1.9 when the dependent variable is replaced with *Ln(1+PATS)* indicate that the effect of our mechanisms stems from its impact on R&D quality rather than on higher patent propensities.

Table 1.10: Innovation and options volume – Profitability

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Ln(Optvol)		-0.308***		-0.127
x Δ ROA$_{t-1}$		(0.116)		(0.082)
Ln(Optvol)	0.176***	0.173***	0.165***	0.163***
	(0.033)	(0.033)	(0.028)	(0.028)
Δ ROA$_{t-1}$	-0.256	0.517	-0.062	0.258
	(0.229)	(0.386)	(0.151)	(0.265)
InstOwn	-0.124	-0.124	-0.245	-0.245
	(0.175)	(0.175)	(0.158)	(0.158)
Ln(K/L)	0.052	0.053	0.066	0.066
	(0.058)	(0.058)	(0.051)	(0.051)
Ln(Sales)	0.121***	0.119**	0.128***	0.127***
	(0.047)	(0.047)	(0.039)	(0.039)
Ln(Age)	-0.088	-0.088	-0.032	-0.032
	(0.094)	(0.095)	(0.081)	(0.081)
Ln(R&D stock)	0.259***	0.260***	0.205***	0.206***
	(0.052)	(0.052)	(0.049)	(0.049)
Observations	2658	2658	2658	2658

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*), lagged change in profitability (Δ ROA_{t-1}), their interaction and other firm-level control variables. Firms in columns: 526. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.5.5 CEO compensation

Up to this point, we have not considered how managerial compensation can help to motivate innovation. Under the optimal contracting view (Holmström and Tirole, 1993), it is efficient for firms in active option markets to grant their managers more stock- and less cash-based pay, as prices are more informative. At the same time, managerial compensation packages that are closely tied to stock price may decrease risk aversion and motivate the manager to expend effort in long-term intangible assets. In particular, it is common to argue that incentives in the form of stock options prevent managers from taking myopic decisions and provide them with increased incentives to take on risky projects. Consistent with this, Coles, Daniel, and Naveen (2006) find that compensation structures with higher vega incentives (controlling for delta) are associated with riskier investment policy as captured by increased R&D, increased focus and reduced PP&E. Similarly, Francis, Hasan, and Sharma (2011) show that incentives in the form of vested and unvested options have a positive and significant effect on patents and citations. Hence, it is possible that part of the positive effect of options trading activity on innovation might be attributable to contractual incentives. We explore this explicitly by conditioning on executive compensation schemes.

The data on compensation comes from ExecuComp. Following prior literature (e.g., Coles, Daniel, and Naveen, 2006), the primary characteristics of compensation that we consider are CEO delta and CEO option holdings vega. Delta is defined as the dollar change in a CEO's stock and option portfolio for a 1% change in stock price, and measures the CEO's incentives to increase stock price. Vega is the dollar change in a CEO's option holdings for a 1% change in stock return volatility; it measures the risk-taking incentives generated by the CEO's option holdings. These values are calculated using the one-year approximation method of Core and Guay (2002). We also control for CEO cash compensation (salary plus bonus) and CEO tenure as the number of years in office may be associated with different compensation schemes.

In column 1 of Table 1.11, we re-estimate Eq. (1.1) on the sub-sample of firms with non-missing compensation variables. The coefficient on *Ln(Optvol)* is 0.157 and significant at the 1% level. In column 2, we add the compensation variables. The coefficients on *Ln(1+CEO vega)* and *Ln(1+CEO delta)* are positive but insignificant while the key coefficient on options volume continues to be positive but becomes smaller in magnitude (i.e., goes down to 0.142), which represents a decrease of approximately 10% from the estimate in column 1.[32] In columns 3 and 4, we repeat the specifications of the first two columns but use patent counts as dependent variable. We observe a similar pattern for the coefficient on *Ln(Optvol)*, i.e., it continues to be positive and significant but declines by about 13% once the compensation variables are included. Interestingly, we also see that the coefficient on *Ln(1+CEO vega)* turns significant, which means that higher vega implies more innovative outputs (as one might expect). Overall, these findings suggest that managerial compensation schemes capture part of the size effect of options trading on innovation, and are thus part of the story, although the specific channel underlying this mechanisms is rather ambiguous.

However, taking Table 1.11 as a whole, we note that options volume has a robust positive effect on innovation across all the specifications, indicating that the relationship between options trading activity and innovation goes substantially beyond compensation structures. This is what one would expect under the career concern explanation. Specifically, although the design of the compensation contract can overcome some of the disincentives to innovate, it does not shield managers from the reputational effects of failed innovation. As Gillan, Hartzell, and Parrino (2009) show, in 2000, 54% of the firms

[32]The insignificant coefficient on CEO delta is consistent with the result in Fang, Tian, and Tice (2014), indicating that greater pay-performance sensitivity is not associated with more innovation. This finding remains unaltered if we replace CEO delta by the scaled wealth-performance sensitivity measure of Edmans, Gabaix, and Landier (2009). In an identical specification to column 2, the coefficient (standard error) on this variable is 0.037 (0.048) whereas the coefficient (standard error) on options volume is 0.137 (0.045).

in the S&P 500 had no explicit employment agreement with their CEOs (i.e., the CEOs were employed "at will"). The median time horizon of the remaining 45% was 3 years. Hence, because CEOs enter the labour market repeatedly, their payoffs are ultimately not determined by explicit contracts, but by the effect their respective reputation has on their ability to contract in the future.[33]

Table 1.11: Innovation and options volume – CEO compensation

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Ln(Optvol)	0.157***	0.142***	0.157***	0.137***
	(0.043)	(0.045)	(0.035)	(0.037)
Ln(1+CEO vega)		0.070		0.106***
		(0.044)		(0.033)
Ln(1+CEO delta)		0.021		0.015
		(0.058)		(0.048)
Ln(1+CEO tenure)		-0.037		-0.032
		(0.050)		(0.042)
Ln(1+CEO cash compensation)		-0.066		-0.094**
		(0.047)		(0.040)
InstOwn	-0.107	-0.125	-0.135	-0.168
	(0.204)	(0.210)	(0.188)	(0.193)
Ln(K/L)	0.013	0.022	0.048	0.058
	(0.085)	(0.086)	(0.070)	(0.069)
Ln(Sales)	0.301***	0.292***	0.254***	0.247***
	(0.072)	(0.073)	(0.063)	(0.062)
Ln(Age)	-0.164	-0.168	-0.075	-0.084
	(0.127)	(0.128)	(0.116)	(0.117)
Ln(R&D stock)	0.148***	0.145***	0.113**	0.106**
	(0.055)	(0.055)	(0.052)	(0.051)
Observations	1845	1845	1845	1845

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume *(Optvol)*, CEO compensation variables and other firm-level control variables. Compensation variables are based on data from ExecuComp over the period 1996 – 2004. *CEO vega* is the dollar change in the CEO's wealth for a 0.01 change in standard deviation of returns; *CEO delta* is the dollar change in the CEO's wealth for a 0.01 change in stock price; Vega and delta values are calculated using the one-year approximation method of Core and Guay (2002). *CEO cash compensation* is the sum of CEO salary and bonus and *CEO tenure* is the number of years the CEO has held the position. Firms in columns: 323. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[33]Incorporating the choice of compensation contracts as a consequence of options trading into our analysis clearly goes beyond the scope of this paper, but endogenizing this decision represents an interesting avenue for future research.

1.6 Discussion and Conclusion

How do financial derivatives affect managerial investment decisions in the real economy? Specifically, do they hinder or promote innovation? This paper attempts to answer these questions by studying the relationship between innovation and options markets.

Our findings contrast with the view that developed financial markets exacerbate myopic behaviour by managers and suggest instead that the presence of informed traders in the options market boosts innovation, even after accounting for R&D investments and the potential endogeneity of options volume. In particular, firms with more options trading activity obtain more patents and patent citations per dollar of R&D invested. We interpret these findings as evidence that the enhanced information efficiency induced by options reduces information asymmetries related to R&D, which provides incentives to managers to invest in innovation. This positive impact could derive from a change in the direction of innovative activities or an increase in R&D spending and productivity. Our findings support the former: firms with greater options trading activity pursue a more creative, diverse and risky innovation strategy.

Our results complement those of Roll, Schwartz, and Subrahmanyam (2009), who find that option markets increase firm valuations by allowing agents to cover more contingencies and by stimulating trading on private information. Specifically, we strengthen their claims by establishing a direct link between options trading activity and managerial investment decisions and show that higher levels of options volume are associated with a more efficient allocation of R&D resources, which then translates into higher firm value. To show this, we rely on two findings. First, Hall, Jaffe, and Trajtenberg (2005) provide evidence that an extra citation per patent boosts a firm's market value by 3%. Second, we repeat the tests in Hall, Jaffe, and Trajtenberg (2005) with a slightly augmented set of control variables. We find that the raw number of patent counts and cite-weighted patent counts have a positive and significant association with a firm's market value.[34] In summary, these pieces of evidence suggest that informed traders in the options market reward successful innovation outcomes with a higher valuation. Together with the core finding of our paper that the main effect of options trading is to alter the quality of innovation outputs, the above evidence seems to reveal one possible "bright" side of financial derivatives: their positive impact on firms' market value by motivating firms to invest in innovative activities.

We discuss several possible mechanisms that could contribute to these findings. First, in line with Aghion, Van Reenen, and Zingales (2013) and contrary to the view that informed traders have a disciplinary effect on managers by "forcing" them to innovate, we find that the presence of informed traders improves the incentives to innovate by reducing

[34]These results are tabulated in Table A.14 in the Appendix.

career concerns. The beneficial effect of options trading is more pronounced when product market competition is intense, when managers are less entrenched, and for younger CEOs. We complement their findings by showing that informed agents play a crucial role in motivating innovation, even if these agents cannot intervene directly in firms' operations (i.e., compared to blockholders). Second, given the pressure from investors to meet profitability targets, decreasing investments in innovation is one of the major real earnings management tools that managers often use to report positive or increasing income. Our analysis indicates that informed traders do recognize the consequences of cutting R&D activities and therefore mitigate myopic investment problems. Lastly, we show that the role of informed trading in motivating innovation exists beyond the structure of managerial compensation and corresponding incentives. Although compensation is a mechanism that links options trading activity and innovation, this effect appears to be substantially dominated by reputation-based incentives, at least in our setting.

While our findings on these mechanisms are consistency with our theory, an unanswered question remains, namely, what is the bottom-line impact of options trading on innovation after accounting for the proposed economic mechanisms. To that end, we directly control for all five mechanism variables and re-estimate an augmented version of Eq. (1.1). The results are tabulated in the Appendix, Table A.15. Overall, we find that options trading continues to be positively and significantly (at the 1% level) related to innovation even after controlling for its dependence on these mechanisms but becomes smaller in magnitude (i.e., goes down from 0.216 to 0.154), which reflects a 29% drop from the baseline model.[35] This suggests that while our mechanisms are able to explain a significant proportion of the positive effect of options trading on innovation, the remaining effect is still strikingly large. Specifically, a coefficient of 0.154 suggests that an increase of 200% in the dollar volume of options traded is associated with a 31% increase in cite-weighted patents. For a median firm in this sub-sample, this implies that an increase in the trading volume from \$15 million to \$45 million leads to approximately 7 additional cite-weighted patents (i.e., from 21 to 28). This is a result of economic as well as statistical significance.

Clearly, we made some simplifying assumptions throughout the paper. Specifically, we assume that managers can influence corporate innovation. However, even in the absence of a specific link, there are several reasons to believe that managers can indeed influence patent-based measures of innovation. As Lerner and Wulf (2007) emphasise, managers can change the compensation schemes of R&D executives towards more long-term incen-

[35]To avoid too many missing values in this test, we fill years missing G-Index with the preceding year's G-Index. Also note that the magnitude of the baseline estimate is different from the one reported in Table 1.2, column 4, because we have only a sub-sample of firms (i.e., firms included in the S&P 1500 index) with non-missing mechanism variable. This allows us to compare the change in coefficients on the same observations.

tives, which can significantly improve the quality of innovative outputs. Managers could also initiate reorganizations with new strategic priorities. For instance, Daniel Vasella, the CEO of Novartis from 1996 to 2010, generated a large increase in R&D productivity with two major strategic moves. First, Vasella expanded Novartis's research from a narrow focus on internal discovery and development capabilities to exploration in new areas through extensive collaborations and the establishment of science-based research institutes. Second, he assigned budget and performance responsibilities over R&D to the business units by setting precise goals, cutting waste, and rewarding successful innovators (Datar and Reavis, 2003).

Moreover, while our study draws on one particular "bright" side of financial derivatives, we are agnostic about how these instruments may affect other stakeholder groups in other ways. Although innovation is important for growth and the wealth of nations, we do not conclude that the greater research productivity shown in our study enhances social welfare. With estimates of the current size of the market for derivatives at approximately $700 trillion, this should, however, be a concern for academics, government regulators, managers, and investors. We leave a proper evaluation of the net effects of financial derivatives for future research.

Chapter 2

Measuring Risk when Expected Losses are Unbounded

2.1 Introduction

Risk measurement is becoming more and more important in Economics, Finance and Insurance. Although the standard deviation has many interesting properties as a risk measure in a Gaussian world, asymmetries and heavy tails imply inconsistencies between the standard deviation and the second order stochastic dominance (or the classical utility functions), and they also make it difficult to interpret the standard deviation in terms of potential capital losses. Several recent approaches have attempted to overcome these drawbacks. In particular, a first line of research deals with the axioms that an index of riskiness must satisfy from an Economic Theory perspective (Aumann and Serrano, 2008), whereas a second line deals with the properties allowing us to interpret a risk measure as potential losses and capital requirements (Artzner, Delbaen, Eber, and Heath, 1999).

This paper focuses on the Artzner, Delbaen, Eber, and Heath (1999) approach with a significant novelty: We allow for risks generating unbounded expected losses. As will be seen, Cauchy or a Pareto distribution presents many mathematical problthe inclusion of risks with unbounded expectation (for instance, risks with a ems when extending the notion of coherent (Artzner, Delbaen, Eber, and Heath, 1999) or expectation bounded (Rockafellar, Uryasev, and Zabarankin, 2006) generalized risk measure, and previous literature has addressed this caveat by losing some desirable mathematical properties. For instance, if we use the value at risk (VaR) as a risk measure then we lose continuity and sub-additivity. Though there are risk measures for heavy tailed risks recovering sub-additivity (or convexity at least, Kupper and Svindland (2011)), we still lose continuity.

This paper overcomes the mathematical problems above by extending a given coherent or expectation bounded risk measure to a limited new setting. For instance, despite the fact that the conditional value at risk ($CVaR$) cannot be continuously extended to the whole space of random risks, we will see that in fact it can be continuously extended to some "smaller" spaces containing some risks with infinite expected value. In practical situations most of the involved risks will have finite expected value, so we should not expect to find infinitely many risks with infinite expectation. More likely, we will just find a few, or even only one. Then, instead of extending a coherent risk measure to a "too large set of risks", we will look for extensions applying if the set of heavy tailed risks is finitely generated.

The outline of the paper is as follows. Section 2 is devoted to introducing notations, the framework, and the main problem we are going to address: The extension of risk measures so as to conserve continuity and sub-additivity (or convexity, at least) and simultaneously involve risks whose fat tail leads to unbounded expected losses. We will summarize the mathematical problems affecting this objective.

Theorem 2 is the main result of Section 3. It states the existence of the required extension if the set of fat tailed risks has a finite generator. Furthermore, Remarks 3 and

4 will show how to construct the extended risk measure in a recursive manner.

Section 4 provides us with illustrative examples and applications. In particular, we will extend the $CVaR$ and the weighted $CVaR$ ($WCVaR$) by "integrating in a coherent manner" these risk measures with the VaR of the heavy tailed risks. We have selected $CVaR$ and $WCVaR$ due to their additional properties, since they are consistent with second order stochastic dominance (Ogryczak and Ruszczynski, 2002) and may be frequently optimized by linear programming methods (Mansini, Ogryczak, and Speranza (2007), see also Konno and Yamamoto (2005)). We will also summarize some actuarial applications, such as some extensions of the expected value premium principle.

The last section of the paper summarizes the most important conclusions.

2.2 Preliminaries and Notations

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ composed of the set of "states of the world" Ω, the $\sigma-$algebra $\mathcal{F}$ and the probability measure $\mathbb{P}$. Denote by $\mathbb{E}(y)$ the mathematical expectation of every $\mathbb{R}-$valued random variable y defined on Ω. Let $1 \leqslant p < \infty$ and denote by L^p the Banach space of random variables y on Ω such that $\mathbb{E}(|y|^p) < \infty$,[1] endowed with the norm

$$\|y\|_p = \left(\mathbb{E}\left(|y|^p\right)\right)^{1/p}.$$

According to the Riesz Representation Theorem, L^q is the dual space of L^p, where $q \in (1, \infty]$ is characterized by $1/p + 1/q = 1$.

Let $[0, T]$ be a time interval. From an intuitive point of view, one can interpret that $y \in L^p$ represents the portfolio pay-off at T for some arbitrary investor (finance), or claims at T for some arbitrary insurer (actuarial science). Throughout this paper y will represent the random wealth at T, although other interpretations would not modify our main conclusions. If

$$\rho : L^p \longrightarrow \mathbb{R}$$

is a risk measure then $\rho(y)$ may be understood as the "risk" associated with the wealth y. Let us assume that ρ satisfies a representation theorem in the line of Artzner, Delbaen, Eber, and Heath (1999) or Rockafellar, Uryasev, and Zabarankin (2006). More precisely, consider the sub-gradient of ρ

$$\Delta_\rho = \left\{z \in L^q; -\mathbb{E}(yz) \leqslant \rho(y), \forall y \in L^p\right\} \subset L^q \tag{2.1}$$

composed of those linear expressions lower than ρ. Δ_ρ will be convex and

[1] and therefore $\mathbb{E}\left(|y|^{p'}\right) < \infty$ if $1 \leqslant p' \leqslant p$. Recall that $L^{p'} \supset L^p$ if $1 \leqslant p' \leqslant p$.

weakly$^*-$compact,[2] and ρ will be its envelope, in the sense that

$$\rho(y) = Max \ \{-\mathbb{E}(yz); \ z \in \Delta_\rho\} \tag{2.2}$$

holds for every $y \in L^p$. Furthermore, we will also assume the existence of $\tilde{E}_\rho \geqslant 0$ such that

$$\Delta_\rho \subset \left\{z \in L^q; \ \mathbb{E}(z) = \tilde{E}_\rho\right\}. \tag{2.3}$$

These assumptions are equivalent to the well-known properties of sub-additivity, homogeneity and translation invariance. To sum up, we have:

Assumption 1. *The risk measure ρ satisfies the equivalent conditions a) and b) below:*

a) The set Δ_ρ given by (2.1) *is convex and weakly$^*-$compact,* (2.2) *holds for every $y \in L^p$, and* (2.3) *holds.*

b) ρ is continuous, sub-additive ($\rho(y_1 + y_2) \leqslant \rho(y_1) + \rho(y_2)$), homogeneous ($\rho(\lambda y) = \lambda\rho(y)$ if $\lambda \geqslant 0$) and $\tilde{E}_\rho-$translation invariant ($\rho(y+k) = \rho(y_1) - \tilde{E}_\rho k$ if $k \in \mathbb{R}$ is zero-variance).

We will not prove the equivalence between Conditions *a*) and *b*) above as similar results may be found in several papers (see, for instance, Balbás, Balbás, and Balbás (2013)).

Assumption 1 is not at all restrictive since it is satisfied by every expectation bounded risk measure with $\tilde{E}_\rho = 1$ and by every deviation measure (Rockafellar, Uryasev, and Zabarankin, 2006) with $\tilde{E}_\rho = 0$. Examples of expectation bounded risk measures are the $CVaR$ and the $WCVaR$, amongst many others. Examples of deviation measures are, amongst others, the classical $p-$deviation

$$\sigma_p(y) = \left[\mathbb{E}\left(|\mathbb{E}(y) - y|^p\right)\right]^{1/p},$$

or the upside and downside $p-$semi-deviations

$$\sigma_p^+(y) = \left[\mathbb{E}\left(|Max \ \{y - \mathbb{E}(y), 0\}|^p\right)\right]^{1/p}$$

and

$$\sigma_p^-(y) = \left[\mathbb{E}\left(|Max \ \{\mathbb{E}(y) - y, 0\}|^p\right)\right]^{1/p}.$$

If $\tilde{E}_\rho = 1$ then it is easy to see that ρ is also coherent in the sense of Artzner, Delbaen, Eber, and Heath (1999) if and only if

$$\Delta_\rho \subset L_+^q = \{z \in L^q; \mathbb{P}(z \geqslant 0) = 1\}.$$

[2] See Rudin [15] for further details about *weakly*$^*-$compact sets in Banach spaces.

Assumption 1 may be relaxed and the main conclusions of this paper remain true. For instance, sub-additivity and homogeneity may be replaced by convexity (in the line of Balbás, Balbás, and Balbás (2010) or Föllmer and Schied (2002)) and $\tilde{E}_\rho-$translation invariance may be removed in 1b). Nevertheless, we prefer to impose Assumption 1 as it is because it significantly simplifies the exposition.

We will also deal with the (metric but not Banach) space L^0. Every random variable belongs to L^0, whose usual metric is given by

$$d\left(y_1, y_2\right) = \mathbb{E}\left(Min\ \left\{\left|y_1 - y_2\right|, 1\right\}\right).$$

It is known that Metric d above leads to the "convergence in probability", which is strictly weaker than the L^p-convergence. As said above, d cannot be given by a norm, and L^0 is not a Banach space. Therefore, the dual space of L^0 may be "too small", and this dual actually reduces to zero if $\mathbb{P}$ is atomless (Rudin [15]). In particular, if a function $\rho : L^0 \longrightarrow \mathbb{R}$ satisfies Condition 1a), then (2.2) implies that $\rho = 0$. In other words:

Remark 1 *Assumption* 1 *cannot be imposed for functionals* $\rho : L^0 \longrightarrow \mathbb{R}$*, because it would imply that* $\rho = 0$ *if* $\mathbb{P}$ *were atomless.*

The latter remark implies that there are no non-null, continuous, sub-additive and homogeneous functionals on L^0 (or even on much smaller proper subspaces of L^0, Delbaen (2009)). Yet in finance, operational risk and insurance one can find risks whose distribution does not belong to L^1, *i.e.*, it does not have a finite expectation (Chavez-Demoulin, Embrechts, and Nešlehová, 2006). Several authors have proposed to use VaR if tails are so heavy that it is impossible to find sub-additive risk measures (Chavez-Demoulin, Embrechts, and Nešlehová (2006), Embrechts, Puccetti, Rüschendorf, Wang, and Beleraj (2014), etc.). Others have studied non-continuous sub-additive risk measures (Kupper and Svindland, 2011). On the other hand, using continuous sub-additive risk measures has many important analytical advantages, since the optimization of such functions is much simpler and many classical financial and actuarial problems (pricing and hedging, portfolio choice, equilibrium, optimal reinsurance, etc.) become much easier to tackle (Balbás, Balbás, and Balbás (2010), among others). For these reasons, it may be worthwhile to look for partial solutions overcoming Remark 1 above, while still preserving some kind of continuity and sub-additivity. This is the main purpose of this paper.

Consider a finite collection of linearly independent final wealths

$$\{w_1, w_2, ..., w_m\} \subset L^0$$

and suppose that their tails are so fat that $w_i \notin L^1$, $i = 1, 2, ..., m$ (i.e., $\mathbb{E}\left(\left|w_i\right|\right) = \infty$, $i = 1, 2, ..., m$). Consider the linear manifold L generated by $\{w_1, w_2, ..., w_m\}$ and suppose that it does not contain non-null elements of L^p. Since L has finite dimension, it only has a unique separated vector topology (Rudin, 1973), and this is the one induced by the topology of L^0. In other words, the sequence $(\sum_{i=1}^{m} x_{i,n} w_i)_{n=1}^{\infty}$ converges in probability to $\sum_{i=1}^{m} x_i w_i$ if and only if $(x_{i,n})_{n=1}^{\infty}$ converges to x_i, $i = 1, 2, ..., m$. Thus, manifold L recovers the structure of a Banach space and we can define non-trivial risk measures on L, that we will denote ρ_L.

Assumption 2. *A risk measure $\rho_L : L \longrightarrow \mathbb{R}$ satisfies the equivalent conditions a) and b) below:*

a) The set

$$\Delta_L = \left\{ (\xi_i)_{i=1}^{m} \in \mathbb{R}^m; \ -\sum_{i=1}^{m} x_i \xi_i \leqslant \rho_L \left(\sum_{i=1}^{m} x_i w_i \right) \ \forall \, (x_i)_{i=1}^{m} \in \mathbb{R}^m \right\} \subset \mathbb{R}^m$$

is convex and compact, and

$$\rho_L (w) = Max \left\{ -\sum_{i=1}^{m} x_i \xi_i; \ \ (\xi_1, \xi_2, ..., \xi_m) \in \Delta_L \right\}$$

holds for every $w = \sum_{i=1}^{m} x_i w_i \in L$.

b) ρ_L is continuous, sub-additive, and homogeneous.

2.3 Extending the risk measure

As said above, L does not present the drawbacks of L^0, and the risk measure ρ_L does satisfy the required properties. According to Remark 1, ρ_L cannot be extended to the whole space L^0 unless we lose its good properties. Thus, let us propose a partial extension that allows the "integration" those risks included in L and those included in L^p. In practical applications we do not expect to find infinitely many risks involving infinite expectations. More likely, we will just find a few (or even only one). Then, the proposed solution may be sufficient, since we will be able to have a "global risk measure" containing both ρ and ρ_L.

In order to jointly manage the risk given by ρ and ρ_L, we need to deal with the space

$$L^p + L = \left\{ y + \sum_{i=1}^{m} x_i w_i \in L^0; \ y \in L^p, \ \ (x_1, x_2, ..., x_m) \in \mathbb{R}^m \right\}$$

which contains those risks included in L^p, those ones included in L, and their linear combinations.

Theorem 2 *There exists an extension $\tilde{\rho}_L : L^p + L \longrightarrow \mathbb{R}$ such that:*

a) $\tilde{\rho}_L$ *is continuous, sub-additive, homogeneous, and* $\tilde{E}_\rho-$*translation invariant (i.e.,* $\tilde{\rho}_L(y+w+k) = \tilde{\rho}_L(y+w) - \tilde{E}_\rho k$ *if* $y \in L^p$, $w \in L$, *and* $k \in \mathbb{R}$*).*

b) $\tilde{\rho}_L(y) = \rho(y)$ *if* $y \in L^p$ *and* $\tilde{\rho}_L(w) = \rho_L(w)$ *if* $w \in L$.

c) $\tilde{\rho}_L$ *is minimal among the functionals* $\Gamma : L^p + L \longrightarrow \mathbb{R}$ *satisfying a) and b).*

d) The set (sub-gradient of $\tilde{\rho}_L$*)*

$$\begin{aligned}&\tilde{\Delta}_L = \\ &\left\{(z,\xi); - \mathbb{E}(yz) - \textstyle\sum_{i=1}^m x_i\xi_i \leqslant \tilde{\rho}_L\left(y + \sum_{i=1}^m x_i w_i\right) \ \forall (y,x) \in L^p \times \mathbb{R}^m\right\} \\ &\subset L^q \times \mathbb{R}^m\end{aligned} \tag{2.4}$$

is convex and weakly$^*-$*compact.*

e)

$$\tilde{\rho}_L(y+w) = Max\left\{- \mathbb{E}(yz) - \sum_{i=1}^m x_i\xi_i; \ \ (z,\xi_1,\xi_2,...,\xi_m) \in \tilde{\Delta}_L\right\} \tag{2.5}$$

holds for every $y \in L^p$ *and every* $w = \sum_{i=1}^m x_i w_i \in L$.

We will skip the proof of Theorem 2 since it is quite technical and beyond the scope of this paper.[3] However, from an intuitive viewpoint, Theorem 2 has a simple interpretation. One can extend both ρ and ρ_L in such a manner that they become "integrated" in a global measure $\tilde{\rho}_L$ which preserves the required properties, despite the fact that $L^p + L$ has infinitely many dimensions and the convergence in this global space still involves convergence in probability. Therefore, one can simultaneously deal with those standard risks y with finite expectations and those much heavier tailed risks w whose expectations are not finite.

Theorem 2 is an existence result, but it does not indicate how to construct $\tilde{\rho}_L$ in practice. Let us address this point.

Remark 3 *(Building* $\tilde{\rho}_L$ *in practice for a single heavy tailed risk). Suppose firstly that* $m = 1$, *i.e.,* L *is a linear manifold generated by only one heavy tailed risk* w *with no finite expected value.*

Step_1. Construct the sub-gradient $\tilde{\Delta}_L$ *of* $\tilde{\rho}_L$ *in such a way that the natural projections* $\Pi_p : L^p + L \longrightarrow L^p$ *and* $\Pi_L : L^p + L \longrightarrow L$ *satisfy* $\Pi_p\left(\tilde{\Delta}_L\right) = \Delta_\rho$ *and* $\Pi_L\left(\tilde{\Delta}_L\right) = \Delta_L$. *It may be easily done as follows:*

Step_2. Fix $\phi \in L^p$,

$$-\rho(\phi) = Min\left\{\mathbb{E}(\phi z); z \in \Delta_\rho\right\} \tag{2.6}$$

[3] The proof uses the Hausdorff maximal principle or, equivalently, the Zorn's lemma (Rudin, 1973)). One must construct familes of *weakly*$^*-$compact sub-sets of $L^q \times \mathbb{R}^m$ whose projections equal Δ_ρ and Δ_L respectively. The intersection of a totally ordered sub-family is non void and minimal for the sub-family, and the Zorn´s lemma leads to the sub-gradient $\tilde{\Delta}_L$ of $\tilde{\rho}_L$.

and

$$\rho(-\phi) = Max\{\mathbb{E}(\phi z); z \in \Delta_\rho\} \tag{2.7}$$

(see (2.2)). Choose ϕ *in such a manner that* $-\rho(\phi) < \rho(-\phi)$ *holds.*[4] *Notice that (2.6) and (2.7) obviously imply that*

$$-\rho(\phi) \leqslant \mathbb{E}(\phi z) \leqslant \rho(-\phi) \tag{2.8}$$

holds for every $z \in \Delta_\rho$.

Step_3. Transform the interval

$$[-\rho(\phi), \rho(-\phi)]$$

into the interval

$$[-\rho_L(w), \rho_L(-w)]$$

by means of the (one to one, unless $-\rho_L(w) = \rho_L(-w)$*) increasing affine function*[5]

$$[-\rho(\phi), \rho(-\phi)] \ni t \longrightarrow F(t) =$$

$$-\rho_L(w) + \frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}(t + \rho(\phi)) \in [-\rho_L(w), \rho_L(-w)].$$

Step_4. $\tilde{\Delta}_L$ *will be chosen according to the affine function above. More precisely,*

$$\begin{aligned}&\tilde{\Delta}_L = \\ &\{(z,\xi); z \in \Delta_\rho \text{ and } \xi = F(\mathbb{E}(\phi z))\} \\ &\subset L^q \times \mathbb{R}.\end{aligned} \tag{2.9}$$

According to (2.8), the construction of $\tilde{\Delta}_L$ *is correct and its weakly*$^*-$*compactness follows from the weakly*$^*-$*compactness of* Δ_ρ *and the continuity of* F *and* $\mathbb{E}$. *It may be also proved that* $\Pi_p\left(\tilde{\Delta}_L\right) = \Delta_\rho$ *and* $\Pi_L\left(\tilde{\Delta}_L\right) = \Delta_L$ *hold, though this proof is complex and therefore omitted.*

Step_5. Once $\tilde{\Delta}_L$ *is known,* $\tilde{\rho}_L$ *is given by (2.5), which becomes now*

$$\tilde{\rho}_L(y + xw) = Max\{-\mathbb{E}(yz) - xF(\mathbb{E}(\phi z));\ z \in \Delta_\rho\} \tag{2.10}$$

---

[4] Notice that $0 = \rho(0) \leqslant \rho(\phi) + \rho(-\phi)$ leads to $-\rho(\phi) \leqslant \rho(-\phi)$. Moreover, the existence of ϕ satisfying the strict inequality holds if ρ is not linear (Δ_ρ is not a singleton), which is the case for all of the usual risk measures.

[5] Once again, $0 = \rho_L(0) \leqslant \rho_L(w) + \rho_L(-w)$ leads to $-\rho_L(w) \leqslant \rho_L(-w)$.

for every $y \in L^p$ and every $x \in \mathbb{R}$. Manipulating,

$$\mathbb{E}(yz) + xF(\mathbb{E}(\phi z)) =$$

$$\mathbb{E}(yz) + x\left(-\rho_L(w) + \frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}(\mathbb{E}(\phi z) + \rho(\phi))\right) =$$

$$\mathbb{E}\left(z\left(y + x\left(\frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}\phi\right)\right)\right)$$

$$+x\left(-\rho_L(w) + \frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}\rho(\phi)\right),$$

and (2.2) and (2.10) imply that

$$\begin{aligned}&\tilde{\rho}_L(y + xw) = \\ &\rho\left(y + x\left(\frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}\phi\right)\right) \\ &+\left(\rho_L(w) - \frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}\rho(\phi)\right)x.\end{aligned} \tag{2.11}$$

Notice that (2.11) leads to $\tilde{\rho}_L(y) = \rho(y)$ if $x = 0$, and $\tilde{\rho}_L(w) = \rho_L(w)$ if $y = 0$ and $x = 1$. In other words, $\tilde{\rho}_L$ really extends ρ and ρ_L.

The selection of ϕ (*Step_2*) is only constrained by the inequality $-\rho(\phi) < \rho(-\phi)$, which generates many degrees of freedom. In other words, we really have a choice when computing in practice the extension $\tilde{\rho}_L$ of Theorem 2, since it is not unique. Next section gives some rules to select ϕ, mainly related to the specific risk $\tilde{\rho}_L(y + xw)$ that one would like to associate with some reachable strategies $y + xw$ of $L^p + L$.

Remark 4 *(Building $\tilde{\rho}_L$ in practice: General case). Let us apply the induction method on the number m of heavy tailed risks. Suppose that we have an extension $\tilde{\rho}_L$ of ρ and ρ_L on $L^p + L_{m-1}$, L_{m-1} denoting the linear manifold generated by $\{w_1, w_2, ..., w_{m-1}\}$. In such a case we have to extend $\tilde{\rho}_L$ to one more dimension, and it is obvious that the methodology described in Step_1 – Step_5 above applies again. Thus, bearing in mind (2.11), we can select $\phi_m \in L^p$ with $-\rho(\phi_m) < \rho(-\phi_m)$ and the global risk measure $\tilde{\rho}_L$ will be given by*

$$\begin{aligned}&\tilde{\rho}_L\left(y + \textstyle\sum_{i=1}^m x_i w_i\right) = \\ &\tilde{\rho}_L\left(y + x_m\left(\frac{\rho_L(-w_m) + \rho_L(w_m)}{\rho(-\phi_m) + \rho(\phi_m)}\phi_m\right) + \textstyle\sum_{i=1}^{m-1} x_i w_i\right) \\ &+\left(\rho_L(w_m) - \frac{\rho_L(-w_m) + \rho_L(w_m)}{\rho(-\phi_m) + \rho(\phi_m)}\rho(\phi_m)\right)x_m.\end{aligned} \tag{2.12}$$

2.4 Examples

Expressions (2.11) and (2.12) provide us with the continuous, sub-additive, homogeneous and $\tilde{E}_\rho-$translation invariant extensions we were looking for. Moreover, they may be easily computed in practice because we have closed formulas, and they are easily optimized (minimized) because we have a convex and *weakly*$^*-$compact sub-gradient (see (2.4), (2.5), (2.9) and (2.10)). Let us show two examples whose unique objective is to illustrate how $\tilde{\rho}_L$ is in practice.

Example 1. Suppose that we would like to use the conditional value at risk $\rho = CVaR_\mu$ with the level of confidence $0 < \mu < 1$, but one of the risks w has a Pareto distribution with unbounded expectation (for instance, this situation may be usual in operational risk, Chavez-Demoulin, Embrechts, and Nešlehová (2006)). Obviously $\rho(w) = CVaR_\mu(w) = \infty$, which means that $CVaR_\mu$ is not continuous any more and, therefore, the computation and optimization of risks such as $CVaR_\mu(y + xw)$, with $y \in L^1$ and $x \in \mathbb{R}$, is quite difficult to address from a mathematical perspective (Balbás *et al* [3]). Then, following Embrechts, Frey, and McNeil (2005), one can deal with the value at risk in order to measure the risk of the heavy tailed distribution w, and then to integrate this risk measure with the conditional value at risk. In such a case, (2.11) leads to the risk measure

$$\begin{aligned} &\tilde{C}VaR_\mu(y + xw) = \\ &CVaR_\mu\left(y + x\left(\frac{VaR_\mu(-w) + VaR_\mu(w)}{CVaR_\mu(-\phi) + CVaR_\mu(\phi)}\phi\right)\right) \\ &+ \left(VaR_\mu(w) - \frac{VaR_\mu(-w) + VaR_\mu(w)}{CVaR_\mu(-\phi) + CVaR_\mu(\phi)}CVaR_\mu(\phi)\right)x, \end{aligned} \tag{2.13}$$

for $y \in L^1$ and $x \in \mathbb{R}$. Expression (2.13) provides us with a continuous, sub-additive, homogeneous and $1-$translation invariant risk measure that extends $CVaR_\mu$ and applies to the Pareto distribution w and its linear combinations with risks of bounded expectation. Furthermore, according to Remark 3, (2.4) and (2.5), and bearing in mind that the $CVaR_\mu-$sub-gradient is (Rockafellar, Uryasev, and Zabarankin, 2006)

$$\left\{z \in L^1; \mathbb{E}(z) = 1,\ 0 \leqslant z \leqslant \frac{1}{1-\mu}\right\},$$

(2.13) may be represented by its convex and *weakly*$^*-$compact sub-gradient, composed of those couples $(z,\xi) \in L^1 \times \mathbb{R}$ such that

$$\begin{cases} \mathbb{E}(z) = 1 \\ 0 \leqslant z \leqslant \frac{1}{1-\mu} \\ \xi = -VaR_\mu(w) + \dfrac{VaR_\mu(-w) + VaR_\mu(w)}{CVaR_\mu(-\phi) + CVaR_\mu(\phi)} \left(\mathbb{E}(\phi z) + CVaR_\mu(\phi)\right) \end{cases} . \tag{2.14}$$

As already said, we have a choice for $\phi \in L^1$, and the selection of this risk affects the final extension $\tilde{C}VaR_\mu$ in (2.13). In practice ϕ may be chosen in such a manner that the risk of some selected heavy tailed distributions still matches their value at risk, since VaR_μ has a nice economic interpretation in terms of potential capital losses. Formally, one can select a collection of risks $\{y_1, y_2, ..., y_k\} \subset L^1$ with bounded expectation, and then choose $\phi \in L^1$ so as to satisfy

$$\tilde{C}VaR_\mu(y_i + w) = VaR_\mu(y_i + w),$$

$i = 1, 2, ..., k$. □

Example 2. According to Remark 4, Example 1 may be easily extended for more than one Pareto (or other heavy tailed) distribution. In other words, if $\{w_1, w_2, ..., w_m\}$ are independent Pareto distributions whose non-trivial linear combinations have unbounded expectations, then one can construct a continuous, sub-additive, homogeneous and $1-$translation invariant $\tilde{C}VaR_\mu$ such that $\tilde{C}VaR_\mu(y) = CVaR_\mu(y)$ if the expectation of y is bounded, and $\tilde{C}VaR_\mu(w_i) = VaR_\mu(w_i)$, $i = 1, 2, ..., m$. Expression (2.12) provides us with the effective construction in a recursive manner. Besides, the role of $CVaR$ may be played by many other coherent risk measures (Wang measure, dual power transform, risk measures given by concave distortions, weighted $CVaR$, etc.), and the role of VaR may be played by alternative selections of every risk $\rho_L(w_i)$ (*ad hoc* selections based on expert opinions and such that $-\rho_L(w_i) \leqslant \rho_L(-w_i)$, VaR with much higher level of confidence than μ if the tails are too fat, etc.). Finally, if the role of the coherent risk measure is replaced by a deviation measure (absolute deviation, standard deviation, semi-deviations, etc.), then we will be creating its continuous, sub-additive, homogeneous and $0-$translation invariant extensions (deviations) that also apply for some risks with unbounded expectation (and therefore unbounded first order central moment, unbounded second order central moment, etc.) and their linear combinations with risks with finite mean value. □

Example 3. (Actuarial application: Extending the expected value premium principle). Many classical financial and actuarial problems (portfolio choice, optimal reinsurance, operational risk, etc.) have been revisited with coherent risk measures. Our extension permits us to involve risks with unbounded expectation.

Let us deal with a particular application whose purpose is just illustrative. Consider a random wealth w whose expectation equals infinite. w endows $\mathbb{R}$ with a probability measure $\mathbb{P}$ on its Borel $\sigma-$algebra $\mathcal{F}$. Obviously,

$$\mathbb{E}\left(|w|\right) = \int_{-\infty}^{\infty} |u| \, d\mathbb{P}\left(u\right) = \infty, \tag{2.15}$$

and therefore $\mathbb{E}\left(w\right)$ and $CVaR_{\mu}\left(w\right)$ do not exist. According to Example 1, we can consider the risk measure $\tilde{C}VaR_{\mu}$, which satisfies $\tilde{C}VaR_{\mu}\left(w\right) = VaR_{\mu}\left(w\right)$.

There are many derivatives $f\left(w\right)$ of w belonging to L^1, and therefore the expected value premium principle $(EVPP)$ applies for them all. If $\lambda \geqslant 0$ is the loading rate, then the price of $f\left(w\right)$ will be given by

$$EVPP_{\lambda}\left(f\left(w\right)\right) = \left(1+\lambda\right)\mathbb{E}\left(f\left(w\right)\right) = \left(1+\lambda\right)\int_{-\infty}^{\infty} f\left(u\right) d\mathbb{P}\left(u\right). \tag{2.16}$$

Interesting particular cases for $f\left(w\right)$ may be the "call-spreads with thresholds $a < b$", given by

$$f\left(w\right) = \begin{cases} 0, \ if \ w < a \\ w-a, \ if \ a \leqslant w \leqslant b \\ b-a, \ if \ w > b \end{cases}$$

but there are much more examples. Since (2.16) applies to price every $f\left(w\right) \in L^1$, we can use the pricing method of Balbás *et al* [5] in order to overcome the caveat implied by (2.15), and therefore we can extend (2.16) and create the pricing rule $EVPP_{(\lambda,\mu,\phi)}$ for risks being linear combinations of w and L^1. Actually, $EVPP_{(\lambda,\mu,\phi)}\left(w\right)$ will be the optimal value of

$$\begin{cases} Min \ \ (1+\lambda)\left[\tilde{C}VaR_{\mu}\left(y-w\right) + \mathbb{E}\left(y\right)\right] \\ \\ y = f\left(w\right) \in L^1 \end{cases}$$

y being the decision variable. According to Balbás *et al.* [5], $EVPP_{\lambda}$ will be extended to $EVPP_{(\lambda,\mu,\phi)}$ in such a manner that it is still continuous on $L^1 + \{xw; x \in \mathbb{R}\}$, sub-additive and homogeneous. Moreover, bearing in mind (2.9), (2.14), and the duality results of Balbás, Balbás, and Garrido (2010), it may be proved that

$$EVPP_{(\lambda,\mu,\phi)}\left(w\right) =$$

$$(1+\lambda)\left[-VaR_{\mu}\left(w\right) + \frac{VaR_{\mu}\left(-w\right) + VaR_{\mu}\left(w\right)}{CVaR_{\mu}\left(-\phi\right) + CVaR_{\mu}\left(\phi\right)}\left(\mathbb{E}\left(\phi\right) + CVaR_{\mu}\left(\phi\right)\right)\right].$$

$\square$

2.5 Conclusions

This paper has proposed a constructive way so as to extend risk measures beyond L^1 and simultaneously preserve good mathematical properties such as continuity, sub-additivity, homogeneity and translation invariance. This may be very useful in order to integrate in a global framework risks with unbounded expectation (for instance, some Pareto distributions) with more standard risks. The good properties of the extended risk measure have favorable implications, in the sense that many classical actuarial and financial problems (pricing, hedging, portfolio selection, equilibrium, optimal reinsurance, operational risk, etc.) may be revisit when some expectations are infinite. Illustrative practical examples have been presented, such as extensions for the conditional value at risk or for the expected value premium principle.

Chapter 3

Modelling Electricity Swaps with Stochastic Forward Premium Models

3.1 Introduction

A growing number of electricity market participants use derivatives contracts as a way of trading synthetic electricity generation plants[1]. Therefore, electricity derivatives are becoming progressively an important part of the global energy commodities market. By far, the most liquid derivatives contracts in the electricity markets are forwards, futures and swaps[2]. Thus, to develop proper models of these electricity contracts is an important step for computing risk measures and for the pricing of other derivatives (e.g. options).

The first pricing models for electricity forwards and swaps relied on the classic cost-of-carry pricing model (i.e. spot-price-based approach). This approach has been reasonably successful in the case of many storable commodities; see Schwartz (1997). In this vein, Hilliard and Reis (1998), Schwartz and Smith (2000), Lucia and Schwartz (2002), Cartea and Figueroa (2005) and Casassus and Collin-Dufresne (2005), among others, presented extensions of the basic model. However, in the case of electricity, this approach is problematic. First, electricity swap contracts imply the delivery of a non-storable commodity (electricity) during a period (e.g. twenty-four hours a day during the whole month of January) instead of the point-in-time delivery period that is common in most other commodity forward contracts. The non-storability of electricity makes the standard no-arbitrage pricing approach hard to justify. Second, the framework based on spot prices is unable to incorporate information about the future (e.g. addition of new generation facilities). Third, the theoretical swap prices generated by using these kinds of models are not necessarily consistent with market prices.

If the pricing model based on spot prices is an appropriate framework for the pricing electricity swap contracts, we should expect that the correlation between spot prices and swap market prices should be relatively high and stable over time for most maturities. However, the empirical evidence suggests otherwise, especially when the maturity of the swap contract is longer than one week. In fact, the historical correlation between electricity spot prices and nearby swap prices is neither high nor particularly stable. The implication is that, in the case of electricity, the spot price is unlikely to be a good proxy for most swap prices. Borovkova and Geman (2006a) document that, for the Nord Pool data, the historical correlation (computed by means of a moving window of the past 60 days) between spot prices and the nearby monthly swap contract ranges from 0.65 to

[1]For instance in Europe the EFET (www.efet.org) a group of more than 100 energy trading companies from 27 European countries promotes energy trading throughout Europe and provides templates of many standardized energy derivatives contracts.

[2]In most electricity markets, forward and futures contracts guarantee delivery of the electricity over a period of time (e.g. monthly or yearly contracts) rather than at a fixed future time. As Benth, Cartea, and Kiesel (2008) argue the nature of these contracts are very similar to a swap exchanging a fixed price for floating (spot) electricity price during the defined period. In fact, swap contracts are integrals of traditional forward contracts, with fixed delivery time. Therefore, in this paper, a swap is a futures contract with delivery over a given period.

-0.15. In our EEX data sample, and using the same procedure, the average value of this correlation is 0.34, but it ranges from 0.87 to -0.54. To put these figures into perspective, Alexander (1999) reports that the average correlation between WTI crude oil spot and NYMEX near monthly futures prices is 0.83 and it ranges from 0.65 to 0.95. In the case of oil log returns, the average value of the correlation is 0.92 and the correlation ranges from 0.12 to 0.99. However in the case of the returns of the EEX electricity swap prices, the average value of the correlation is negative (-0.005), and the correlation ranges from 0.45 to -0.35[3].

For these reasons, we propose a pricing approach based on modelling the swap curve directly, and allowing for a very general distribution for the sources of uncertainty (innovations) influencing the state variables. Our model features factors accounting for (a) the average swap price within each market segment, (b) the deterministic seasonal factor and, (c) the stochastic changes in the swap curve shape. The model relies on a particular case of the Multivariate Generalized Hyperbolic (MGH) distributions, namely the Multivariate Normal Inverse Gaussian (MNIG) distributions, which allow stochastic dynamics in terms of correlated NIG Lvy processes. We test the model using data from the German Power Market (European Energy Exchange, EEX) which is Europes largest power market in terms of consumption.

This study extends current literature in several ways. First, Kiesel, Schindlmayr, and Börger (2009) suggest a two-factor model for electricity swaps calibrated to at-the-money options on electricity swaps traded in the EEX market. However, the degree of liquidity of these options varies substantially across contracts and over time. Thus, we prefer working directly with the swap prices, because, at least some of them, are usually highly liquid all the time. Second, in contrast with Benth, Cartea, and Kiesel (2008) who allow for just one Brownian motion as the driver of the dynamics of the swap curve, we argue that correlated Lvy process for each market segment are useful to explain the dynamic behaviour of swap prices, allowing for a more realistic representation. Third, Fleten and Lemming (2003), Audet, Heiskanen, Keppo, and Vehviläinen (2002), Koekebakker and Ollmar (2005) and Bjerksund, Rasmussen, and Stensland (2010) build a continuum of instantaneous-delivery forward contract by smoothing market prices, or by combining market prices with forecasts generated by bottom-up models. However, we argue that working directly with the most liquid swap prices is more transparent, instead of using ad hoc numerical procedures to extract smooth curves from quoted prices[4]. In our case, the liquid contracts are the six contracts closer to maturity for the monthly, quarterly and

[3]See also the evidence in Benth, Benth, and Koekebakker (2008) for the Nord Pool.

[4]Advocates of the smoothing algorithms posit that futures prices with fixed time to maturity can be extracted each day from the smoothed curve. The main criticism is however that the prices are not "true" market prices but interpolations and these smoothed prices may distort the empirical analysis. We prefer to concentrate on actual market prices and include the effect of the changing time to maturity in the SFP component

yearly delivery periods.

Summing up, our contributions are as follows. First, we propose a new stochastic forward premium model that includes exposures to average swap prices within each market segment, deterministic seasonal factors and mean-reverting stochastic deviations from the average price following MNIG processes. This model effectively captures realistic, time-varying characteristics in swap prices, overcoming the limitations of standard models of swap curves that cannot account for asymmetries and fat tails. Second, our empirical results during the period from 2004 to 2013 provide strong evidence supporting our model in comparison with other alternatives, because it presents better empirical fit than spot price-based or traditional HJM-based models. Third, we show that models who do not account for the impact of non-normality are not able to replicate market prices, and, in particular, they are unable of taking account of tail risk. This result is important because González-Pedraz, Moreno, and Peña (2014) present evidence suggesting that tail risk measures for energy portfolios based on standard assumptions (e.g. normality) underestimate actual tail risk, especially for short positions and short time horizons. Another important practical implication of our VaR analysis is that that the capital charges to traders using EEX electricity swap contracts (based on risk adjusted capital under the normality assumption) should be adjusted (most likely upwards) and that the evaluators of the performance of the traders should adjust their recommendations accordingly.

The rest of paper is as follows. In Section 3.2, we briefly revise related literature and present our model. After we describe the data in Section 3.3, we report the results of the empirical analysis in Section 3.4. Section 3.5 compares the model against alternative approaches. Section 3.6 presents some Value-at-Risk results. Section 3.7 concludes.

3.2 The Model

In this section, we outline the basic characteristics of the theoretical model and present some guidelines for its practical implementation. Given that we make use of the MNIG distribution, we suggest readers Embrechts, Frey, and McNeil (2005) for definition and properties of this distribution.

3.2.1 Literature Review

The models for pricing and hedging financial derivatives on energy prices belong to three broad categories: (i) based on fundamental equilibrium (ii) based on spot energy prices and other key variables such as convenience yields or interest rates and (iii) based on forward price processes. Models in the first category focus on supply and demand relationships to obtain the power prices as a solution of an optimization problem. This optimization problem embodies information on market prices and trading activity. This allows the computation of the forward prices, using the condition that they provide equilibrium in the demand for forward contracts; see Bessembinder and Lemmon (2002). In a similar vein, Supatgiat, Zhang, and Birge (2001) show that market-clearing prices are determined by solving a Nash equilibrium problem for the bidding strategies of market agents. Although useful for a wide range of applications, models in this category do not capture appropriately the price dynamics, which is what market participants need in order to develop effective hedging and risk management strategies.

The second category is based on specifying stochastic processes of the spot price and possibly of a limited set of other state variables, calibrating their parameters using market data. Then, one may resort to closed formulas or numerical approximations, in order to price contingent claims. Examples of this approach are Schwartz (1997), Hilliard and Reis (1998), Schwartz and Smith (2000), and Casassus and Collin-Dufresne (2005), among others. This approach has been successfully applied to some energy commodities, particularly crude oil, but its adequacy is less clear in the case of electricity markets because of the very specific features present in electricity spot prices. These characteristics are strong seasonality, mean reversion, jumps, stochastic volatility and regime switching (see Escribano, Ignacio Peña, and Villaplana (2011) among others) and are caused by the difficulties of storing electricity efficiently; see also Lucia and Schwartz (2002), Geman and Roncoroni (2006) and Cartea and Figueroa (2005). Besides that, this approach has some disadvantages such as its inability to incorporate information about the future (e.g. decreases in supply caused by planned maintenance of generation facilities) and because endogenously generated swap prices are not necessarily consistent with observable market prices. Quinn, Reitzes, and Schumacher (2005) argue that electricity swap prices are a function of market expectations of demand and cost conditions during the actual delivery period, and these expectations are not necessarily influenced by current market behaviour (i.e. spot prices), and they present evidence supporting their claim in the PJM market. Furthermore, the historical correlation between electricity spot prices and the nearby swap prices is not particularly stable, which suggests that the spot price is not a good proxy for the futures prices. Benth and Koekebakker (2008) find that contracts located in the very short end of the swap curve are the only ones with sizeable correlations with the spot price at Nord Pool.

The third category is based on the direct modelling of the term structure of the electricity swap prices. Within this category, a first line of research relies on Heath, Jarrow, and Morton (1992) (HJM) which focus on the dynamics of the swap curve as a whole. Examples of this approach are Cortazar and Schwartz (1994), Amin et al. (1995), Miltersen and Schwartz (1998), Clewlow and Strickland (2000), Koekebakker and Ollmar (2005), Miltersen (2003), Audet, Heiskanen, Keppo, and Vehviläinen (2002) and Trolle and Schwartz (2009), among others. In all these cases, the market swap price curve is an input into the derivative pricing model and therefore derivatives prices thus generated should be consistent with observable swap market prices. A second line of research focuses on modelling a given function of observed swap prices and then analysing stochastic deviations from this function by means of additional state variables. An example of this approach is Borovkova and Geman (2006b), who propose to use a parsimonious two-factor model, in which the first factor is the average forward price and the second factor is analogous to the stochastic convenience yield. However, a potential problem of all the previous models within this category is their assumption of Gaussian distributions for the innovations of the stochastic processes[5]. As suggested by Frestad, Benth, and Koekebakker (2010) in the case of electricity swaps, this assumption is unlikely to be appropriate, because the innovations of the electricity swap prices are strongly non-normal. Besides that, the factor structure of the swap curve in the electricity markets is probably much more complex than in another energy markets. For instance, Koekebakker and Ollmar (2005) report that, in order to explain more than 98% of the variation in the sample covariance matrix in the Nord Pool, they need more than ten factors. Interestingly, factors explaining a large proportion on the variation in the long end of the curve seem to have very low explanatory power in the short end of the curve. This fact suggests that, besides a few general factors affecting the whole curve, some parts of the curve are exposed to unique risk factors that other parts of the curve are not exposed to. Frestad (2008) find strong support for this hypothesis in the Nordic electricity market by using this idea of common and unique factors.

In most papers, the innovation processes driving the state variables are Gaussian. In this paper, we prefer to use a more general process for the innovations. There is growing evidence suggesting that the Normal Inverse Gaussian (NIG) distribution, which was first used for modelling speculative returns in Barndorff-Nielsen and Prause (2001), fits heavy-tailed and skewed financial data well and is, at the same time, analytically tractable, see, for instance, Rydberg (1999), Barndorff-Nielsen and Prause (2001), Forsberg and Bollerslev (2002), and Karlis (2002) among others. Therefore the NIG distribution is especially suitable for the modelling of financial prices, and in particular for the term structure individual contract dynamics (Benth, Cartea, and Kiesel, 2008) as well as their joint evolution (Andresen, Koekebakker, and Westgaard, 2010b). These are the reasons why we choose the MING distribution as our preferred choice.

[5]Exceptions are Andresen, Koekebakker, and Westgaard (2010a) who present a discrete random-field model based on the multivariate NIG distribution, and Di Poto and Fanone (2012) who apply a Lévy multifactor market model by using Independent Component Analysis.

3.2.2 The General Model

Assume $T < \infty$ and let (Ω, F, Q) be a complete filtered probability space, with an increasing and right continuous filtration $[F_t]_{t\in[0,T]}$ where, as usually, F_0 contains all sets of probability zero in F. We assume that the market trades swap contracts with different delivery periods and a bond that yields a constant risk free rate $r > 0$, so futures and forward prices are equal. Consider the price $F_i(t,T)$ of a swap contract with expiry date $T = 1, \dots, N$ which is also the start of the delivery period, time to maturity $\tau = (T - t)$,and delivery length period given by the subscript $i = 1, \dots, I$ (e.g. $i = 1$ (M) for monthly contracts, $i = 2$ (Q) for quarterly contracts and $i = 3$ (Y) for yearly contracts) . These contracts are settled against the daily average spot price during the delivery period and, in agreement with market practice, we call them electricity swaps. Hence, $F_i(t,T)$ denotes the price vector of a completely observable swap curve at time t, where vector T indicates the different swap contracts starting delivery dates which are available at trading date t, and the vector subscript i denotes the underlying asset (delivery period) over which is defined each curve's swap contract. Notice that in a general term structure model, we consider the evolution of the continuum of swap contracts for all the possible expirations. Since only subsets of them trade in the market, we develop a version of the model that specifies the dynamics only for tradable contracts that are liquid enough. Therefore our model follows the spirit of the market model, see Brace, Musiela, et al. (1997), which was originally developed for interest rates. Borovkova and Geman (2006b) propose a similar model to ours, but they consider only two sources of uncertainty, driven by uncorrelated Gaussian innovations. We instead allow for multiple sources of uncertainty, and these sources of uncertainty follow correlated MNIG distributions. In doing, so our model captures the correlation among the sources of uncertainty, as well as some salient stylized facts in the swap electricity market, such as extreme kurtosis. The starting point of our analysis is the following stochastic forward premium model (SFP henceforth), which relates swap prices for any maturity T and delivery length period i as follows,

$$F_i(t,T) = \overline{F}_t exp\left(s_i(K) + \gamma_i(t,T)\right) \tag{3.1}$$

where the first component in the right-hand side is the average level of the swap price within each market segment i defined as the geometric average of the current swap prices as follows

$$\overline{F}_i(t) = \sqrt[N]{\prod_{T=1}^{N} F_i(t,T)} \tag{3.2}$$

and N is the maximum liquid maturity[6]. The average swap price does not contain

[6]For instance if there are three delivery periods, the first with a length of one month, the second with

seasonal factors, which are included in the deterministic seasonal premia factor $s_i(K)$. We define this premium as the collection of long-term average premia on swaps expiring in the calendar period (week, month, quarter) with respect to the average swap price. Depending on the value of i, K can take 52 (weekly), 12 (monthly) or 4 (quarterly) different values. By construction the average of the seasonal factors for each i must be zero. For instance in the case of which $i = M$, that is a monthly (M) delivery period, there are twelve seasonal factors $s_M(K)$, $K = 1, \ldots, 12$ that is, a deterministic collection of 12 parameters.

The quantity $\gamma_i(t,T)$ is the stochastic forward premium (SFP) for the delivery period i and expiry date T. By construction, this SPF is zero on average, and we define it as

$$\gamma_i(t,T) = lnF_i(t,T) - ln\overline{F}_i(t) - s_i(K) \tag{3.3}$$

Next, we specify the stochastic dynamics of the state variables in terms of Multivariate Normal Inverse Gaussian (MNIG) Lévy processes. Notice that the dimension of the system is $d = I + I \times N$. Let $L(t)$ be a d-dimensional vector of MNIG Lévy processes. This means that it has stationary and independent increments, in the sense that the distribution of $L(t) - L(s)$, $t > s \geqslant 0$, is only dependent on $t - s$ and not on t and s separately, such that its increments $dL(t) = L(t + dt) - L(t) = X$ are, standardized (i.e. zero-mean, unit variance) and MNIG distributed with probability density function $MNIG_d\,(X; \alpha, \beta, \delta, \mu, \Sigma)$

$$f(X) = \frac{\delta}{2^{(d-1)/2}} \left[\frac{\alpha}{\pi q(x)}\right]^{(d+1)/2} K_{\frac{d+1}{2}} [\alpha q(x)] \times \exp p(x) \tag{3.4}$$

where,

$$q(x) = \sqrt{\delta^2 + (x-\mu)'\Sigma^{-1}(x-\mu)},\ p(x) = \delta\sqrt{\alpha^2 - \beta'\Sigma\beta} + \beta'(x-\mu)$$

and $K_{\frac{d+1}{2}}$ is the modified Bessel function of the second kind with index $(d+1)/2$, and the parameters have the following characteristics, $\delta > 0$,$\alpha^2 > \beta'\Sigma\beta$, $\beta \in \mathsf{R}^d$, $\mu \in \mathsf{R}^d$, $\Sigma = \sigma_{ij} \in R^{d \times d}$ and we require Σ to be positive definite and $|\Sigma| = 1$. The mean vector of X is

$$E[X] = \mu + \delta\Sigma\beta\sqrt{\alpha^2 - \beta'\Sigma\beta}$$

and the covariance matrix $V[X] = v_{ij}$ $i = 1, \ldots, d$; $j = 1, \ldots, d$ is defined as

a length of three months (one quarter) and the third with a length of twelve months (one year), thus $i = 1$ (M),2(Q),3(Y) and $I = 3$. If there are six liquid maturities 1 to 6, then $T = 1, 2, \ldots, 6$ and $N = 3$. In summary, the liquid contracts are M1, M2, M3, M4, M5, M6, Q1, Q2, Q3, Q4, Q5, Q6 and Y1, Y2, Y3, Y4, Y5 and Y6.

$$V[X] = \delta\sqrt{\alpha^2 - \beta'\Sigma\beta}[\Sigma + (\alpha^2 - \beta'\Sigma\beta)^{-1}\Sigma\beta\beta'\Sigma \tag{3.5}$$

Note that if the skewness parameter is zero $\beta = 0$ the mean vector coincides with μ and the covariance matrix Σ solely determines the correlation structure. For asymmetric MNIG distributions, the correlation structure depends on all parameters, excepting μ.

The marginal distributions of the MNIG distribution are univariate NIG distributions (Lillestol, 2000). Denoting the parameters of the marginal distributions of the i th component of X as X_i and using the obvious notation, the scale-free indicators of asymmetry χ_i and kurtosis ξ_i, $\{(\chi_i, \xi_i) \in \mathsf{R}^2; |\chi_i| < \xi_i < 1\}$ are respectively defined as (Rydberg, 1999):

$$\chi_i = \frac{\xi_i\beta_i}{\alpha_i} \; ; \xi_i = \frac{1}{\sqrt{(1 + \delta_i\gamma_i^2}}$$

where $\gamma_i = \sqrt{\alpha_i^2 - \beta_i^2}$. If the two scale-free parameters are close to zero (i.e. $(\chi_i, \xi_i) \approx (0, 0)$) the marginal NIG distribution is close to being normal. On the other hand, when the scale-free parameter ξ_i is very close to one, the marginal NIG distribution is very similar to the heavy-tailed Cauchy distribution.

We define de dynamics of $ln\overline{F}_i(t)$ and $\gamma_i(t,T)$ under the market probability measure by the stochastic differential equations:

$$dln\overline{F}_i(t) = \kappa_i(\zeta_i - ln\overline{F}_i(t))dt + \theta_{\overline{F}_i} dL_{\overline{F}_i}(t); \text{ i=1,...,I} \tag{3.6}$$

$$d\gamma_i(t,T) = -\omega_{i,T}\gamma_i(t,T)dt + \theta_{\gamma_i(T)} dL_{\gamma_i(T)}(t); \text{ i=1,...,I T=1,...,N} \tag{3.7}$$

The SFPs are subject to their own sources of uncertainty, given by the standardized MNIG Lévy processes, $dL_{\overline{F}_i}(t)$, and $dL_{\gamma_i(T)}(t)$ which are assumed to be correlated. We can substitute (3.6) and (3.7) into (3.1) and derive the dynamics of the swap log-prices under the market probability measure as follows:

$$\begin{aligned} dlnF_i(t,T) = &\left[\kappa_i(\zeta_i - ln\overline{F}_i(t)) - \omega_{i,T}\gamma_i(t,T) + s_i(K)\right] dt \\ &+ \theta_{\overline{F}_i} dL_{\overline{F}_i}(t) + \theta_{\gamma_i(T)} dL_{\gamma_i(T)}(t); \text{ i=1,...,I T=1,...,N} \end{aligned}$$

And therefore $lnF_i(t,T)$ is obtained by integrating the above differential equation with the initial condition

$$lnF_i(0,T) = ln\overline{F}_i(0) + s_i(K) + \gamma_i(0,T)$$

The term structure of swap prices variances is:

$$\varphi_i^2(t,T) = \eta_i^2 + \tau_{i,T}^2 + 2\theta_{\overline{F}_i}\theta_{\gamma_i(T)}\upsilon_{\overline{F}_i,\gamma_i(T)} \tag{3.8}$$

where $\eta_i^2 = \theta_{\overline{F}_i}^2 \times \upsilon_{\overline{F}_i}$ and $\tau_{i,T}^2 = \theta_{\gamma_i(T)}^2 \times \upsilon_{\gamma_i(T)}$ are the variances of the corresponding average factor and of the stochastic discount factor respectively and $\upsilon_{\overline{F}_i,\gamma_i(T)}$ is the covariance between the average factor and the stochastic discount factor, all of them elements of (3.5).

As we are interested in pricing other derivatives, we now consider how to price these derivatives in the risk-neutral world. First, notice that we specify the dynamics (3.1), (3.6) and (3.7) under the market (real-world) probability measure P; therefore, we must select a risk-neutral probability measure Q. A common choice (see Benth, Benth, and Koekebakker (2008)) is the Esscher transform which generalizes the Girsanov transform to Lévy processes and guarantees that the $L(t)$ process is still a NIG Lévy process under Q. This Esscher transform implies that, under the risk-neutral measure Q, the transformed $L^Q(t)$ vector is MNIG-distributed with density function $MNIG_d\left(X; \alpha, \beta + \eta, \delta, \mu, \Sigma\right)$, or in other words, the Esscher transform only changes the asymmetry of the process. The vector μ measures the price of jump risk, that is, the price that market players charge for assuming the risk of not being able to hedge. A positive price leads to a more right-skewed distribution. Given an estimate of the vector μ, and the transformed process under Q, we can apply standard techniques to price other derivatives such as options.

3.2.3 Implementation

Assume we have an historical dataset of n daily swap curves $F_i(t,T)$ $t = 1, \ldots, n$, where vector T indicates the different swap contracts starting delivery dates which are available at trading date t, and the vector subscript i denotes the underlying asset (delivery period) over which is defined each curve's swap contract. Writing (3.1) in logarithm form

$$lnF_i(t,T) = ln\overline{F}_i(t) + s_i(K) + \gamma_i(t,T) \tag{3.9}$$

The least squares optimal estimator for $ln\overline{F}_i(t,T)$ is simply the arithmetic average of log-swap prices within each market segment i. We estimate factor $s_i(K)$ by

$$\widehat{s}_i(K) = \frac{1}{n} \sum_{T_k \in A_i} \sum_{t=1}^{n} \left(lnF_i(t,T_k) - ln\overline{F}_i(t) \right) \tag{3.10}$$

where A_i are the sets of available maturities at time t for contracts with delivery period $i = 1, \ldots, I$, and the index K takes different values k. The seasonal factors must be zero on average for each seasonal period (e.g. monthly) , and then $\sum_{k=1}^{K} s_i(k) = 0$. We estimate the SFP factor by means of the equation

$$\widehat{\gamma}_i(t,T) = lnF_i(t,T) - ln\overline{F}_i(t) - \widehat{s}_i(K) \tag{3.11}$$

For the sake of clarity, we use the notation i=1(M),2(Q),3(Y) and $T = 1, \ldots, 6$ in what follows. Given the complexity of the estimation, we apply a two-step procedure. In the first step we estimate simultaneously all the parameters of mean reversion and volatility in discrete-time versions of Equations (3.6) and (3.7) by means of a system of seemingly unrelated regression equations (SURE). In doing so, we assume that error terms may have cross-equation contemporaneous covariance . The system takes the form

$$\begin{pmatrix} \nabla ln\overline{F}_i(t) \\ \nabla\widehat{\gamma}_i(t,T) \end{pmatrix} = \begin{pmatrix} \kappa_i(\zeta_i - ln\overline{F}_i(t-1)) \\ -\omega_{i,T}\gamma_i(t-1,T) \end{pmatrix} + \begin{pmatrix} \theta_{\overline{F}_i}\epsilon_{\overline{F}_i} \\ \theta_{\gamma_i(T)}\epsilon_{\gamma_i(T)} \end{pmatrix} \tag{3.12}$$

The system contains 21 equations $(3 + 6 \times 3)$ and this SURE model is estimated using the feasible generalized least squares (FGLS) method. Given the dimensions of the problem and that in the second step we require the covariance matrix of the MNIG process Σ to be positive definite and $|\Sigma| = 1$ a convenient normalization suggested in Urzua (1997) is as follows. Let the residuals form Equation (3.12) be defined as

$$Y = \begin{pmatrix} \epsilon_{\overline{F}_i} \\ \epsilon_{\gamma_i(T)} \end{pmatrix}, t = 1, \ldots, n \tag{3.13}$$

where Y (dimensions $d \times n$) has a mean vector of zero and covariance matrix $\Omega = {}_{(i,j)}$. Let Γ denote the orthogonal matrix whose columns are the standardized eigenvectors of Ω, and Λ denote the diagonal matrix of the eigenvalues of Ω . Define $\Omega^{-1/2}$ as the inverse of the square root decomposition of Ω; or, in other words, that

$$\Omega^{-1/2} = \Gamma\Lambda^{-1/2}\Gamma' \tag{3.14}$$

Then the random variable

$$X = \Omega^{-1/2}Y \tag{3.15}$$

has a zero mean vector, and an identity matrix as its covariance matrix. This variable X, which contains the standardized and orthogonal residuals, is then used in the second step.

In the second step we use X as the estimation of the vector $dL(t)$, which is assumed to be MNIG distributed with probability density function $MNIG_d(X; \alpha, \beta, \delta, \mu, \Sigma)$ and we estimate the corresponding parameters of the MNIG distributions by using the EM algorithm developed by Øigård, Hanssen, Hansen, and Godtliebsen (2005) which is an extension of Karlis (2002). We compute bootstrapped confidence intervals based on 500 replications. In order to compute the term structure of swap prices given by (3.8) we set $\varphi_i^2(t,T) = \eta_i^2 + \tau_{i,T}^2 + 2\theta_{\overline{F}_i}\theta_{\gamma_i(T)}\omega_{\overline{F}_i,\gamma_i(T)}$ where $\eta_i^2 = \theta_{\overline{F}_i}^2 \times \omega_{\overline{F}_i}$ and $\tau_{i,T}^2 = \theta_{\gamma_i(T)}^2 \times \omega_{\gamma_i(T)}$ are the variances of the corresponding average factor and of the stochastic discount factor respectively and $\omega_{\overline{F}_i,\gamma_i(T)}$ is the correlation between the two factors, all of them obtained from (3.12).

3.3 Data

Germany is Europe's largest power market and therefore we choose this market as the source of the data. Our data set consists of daily data from June 1, 2004 until December 31, 2012, on settlement prices for the following available baseload swap contracts traded in EEX: Yearly, Quarterly and Monthly (Phelix-Base/Month/Quarter/Year-Futures). The company operating EEX market (EEX AG) has provided the data. We choose the six most liquid contracts within each market segment, that usually are the closest to maturity ones. Within each market segment, these six contracts represent the 99% (100%), 97% (99%) and 100% (100%) of the total trading volume (open interest) in the case of monthly, quarterly and yearly contracts, respectively. We explain the relative liquidity of each contact in more detail in Section 3.2. We define the continuous series as a perpetually linked series of swap settlement prices. For example, M1 starts at the nearest contract month, which forms the first values of the continuous series, until the contract reaches its expiry date, or until the first business day of the actual contract month. At this point, we take the next trading contract month. For all series, we compute the returns as the first difference of log prices.

3.3.1 Summary Statistics

We show the graphs for all swap price series in Figure 3.1 by market segments. The monthly series seem to be the more volatile followed by the quarterly series, being the yearly series the most stable one.

Figure 3.1: Panel A. Patent counts

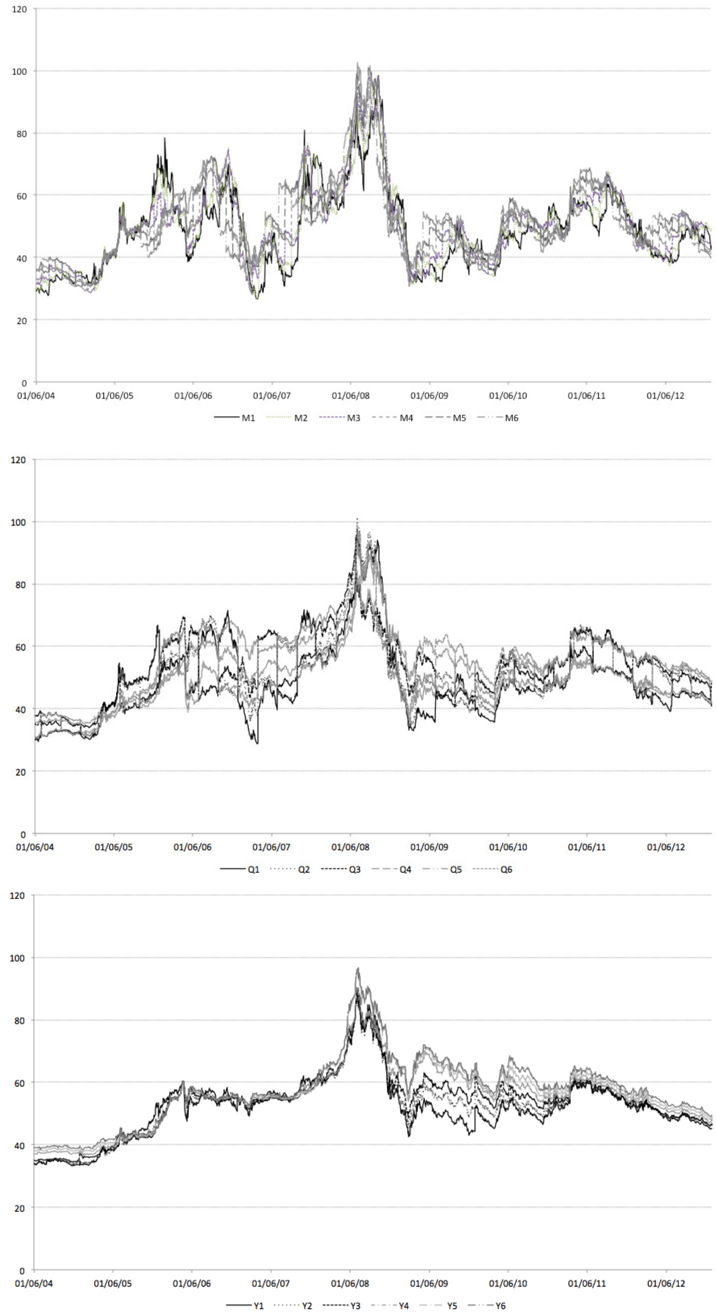


Table 3.1 provides information on the basic statistics for all series in levels. We show the individual series in Panel A and the averages of each market segment in Panel B. Looking at the averages of each market segment (monthly, quarterly and yearly), the average price tends to increase with the maturity of the contract and the volatility tends to decrease with the maturity of the contract, as expected. Volatility is usually higher for the closest-to-maturity contract (Samuelson effect), confirming the well-known fact that short-dated swaps tend to be more volatile than long-dated swaps. Within each class, the average price follows the same pattern, but the volatility has behaviour that is more complex. In the case of monthly contracts, the volatility has an inverted u-shape, in the

case of quarterly contracts, it has a u-shape and in the case of yearly contracts, it increases with time to maturity. In the returns series (not shown) the volatility follows the expected pattern, decreasing with time to maturity and ranging (in annualized terms) from 33% in the case of M1 to 13% in the case of Y6. All series present positive asymmetry and significant kurtosis, suggesting that the normality assumption is not suitable for these series.

Table 3.1: Summary Statistics

	M1	M2	M3	M4	M5	M6	Q1	Q2	Q3	Q4	Q5	Q6	Y1	Y2	Y3	Y4	Y5	Y6
Panel A: Daily Swap Prices																		
Mean	48.74	49.92	50.60	50.99	51.45	51.50	50.44	51.47	51.46	51.74	52.47	52.76	52.00	52.61	53.52	55.23	56.69	57.54
Median	47.88	48.85	49.20	49.17	49.13	49.66	48.75	49.60	49.98	51.79	52.41	51.84	51.94	53.70	54.59	55.30	56.15	57.30
Maximum	98.41	96.76	98.23	101.94	101.00	102.75	97.50	100.93	94.95	84.75	94.07	98.33	90.15	89.00	89.67	90.30	96.30	96.80
Minimum	26.50	26.45	28.25	29.05	30.35	30.55	28.69	31.17	31.05	30.50	30.55	31.70	33.12	33.70	34.40	36.69	37.51	38.29
Std. Dev.	12.82	12.95	13.15	13.01	12.86	12.48	12.76	12.25	10.89	10.33	11.30	11.54	9.90	9.58	9.73	10.04	11.09	11.13
Skewness	0.80	0.79	1.04	1.23	1.25	1.19	0.98	1.19	0.56	0.36	0.72	0.81	0.64	0.30	0.27	0.25	0.58	0.55
Kurtosis	3.76	3.68	4.29	5.06	5.37	5.37	4.25	5.28	3.43	3.21	4.22	4.65	4.72	4.48	4.33	3.76	4.18	4.07
JB	287.2	269.4	541.4	939.8	1072.9	1024.5	488.0	985.3	132.3	51.67	324.9	485.2	415.6	232.4	187.5	73.78	249	215.5
Probability	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Panel B: Average Series M, Q and Y			
	M	Q	Y
Mean	50.53	51.73	54.60
Median	48.98	50.73	54.83
Maximum	99.85	95.09	92.04
Minimum	28.53	30.61	35.62
Std. Dev.	12.88	11.51	10.24

Notes: This table reports some descriptive summary statistics for daily swap prices over the full sample period from 6/1/2004 to 12/31/2012. The sample size is 2179 observations. JB is the Jarque-Bera normality test. Probability is the p-value of this test under the null of normality.

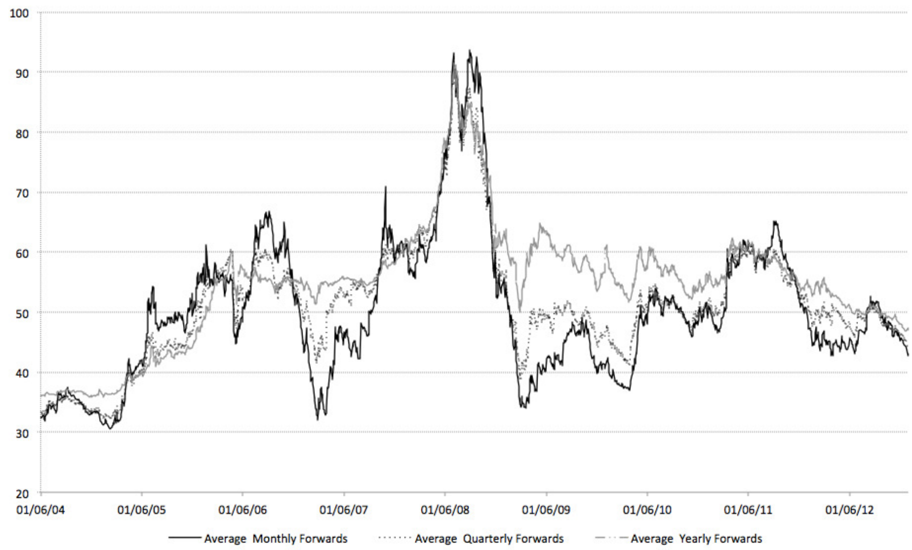
3.4 Empirical Results

In this section, we present the estimation of the components of the swap price as defined in Equation (3.1), that is, the average swap price, the seasonal components and the stochastic forward premium.

3.4.1 Average swap prices

We compute average swap prices as the simple average of swap log prices within each market segment. Figure 3.2 contains the graphs and the basic statistics are in Table 3.2.

Figure 3.2: Average Swap Prices



The basic statistics of the average swap price are consistent with the ones in Table 3.1 in the sense that average price increases with maturity and volatility decreases with maturity. The average return is close to zero, the volatility is higher in the monthly segment and it is lower in the yearly segment, and all series have positive asymmetry and high kurtosis. It is interesting to note that the correlation (both in levels and in first differences) is far from one in the case of the averages in the monthly and yearly segments. This fact stresses the convenience of working with a model based on a specific average component for each market segment.

Table 3.2: Summary Statistics

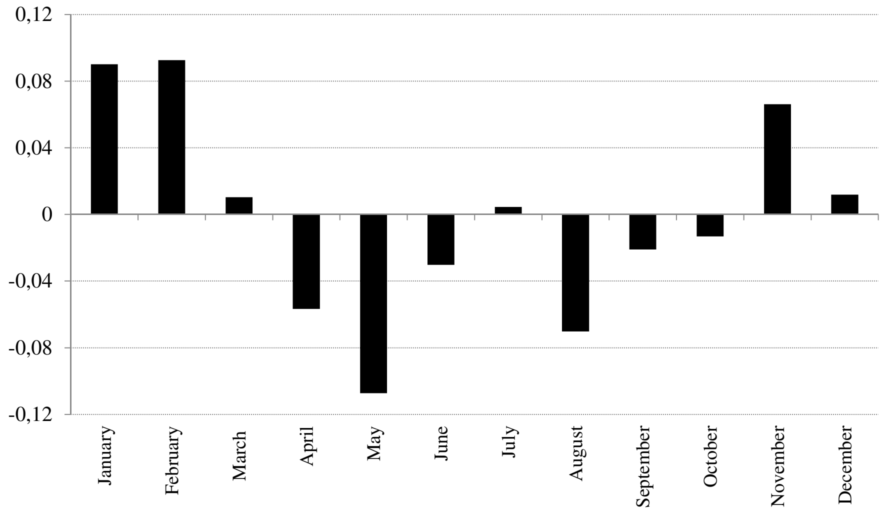
	M	Q	Y
Panel A: $\bar{F}$			
Mean	50.330	51.470	54.520
Median	48.630	50.520	55.190
Maximum	93.630	92.470	91.620
Minimum	30.410	32.210	35.780
Std. Dev.	12.010	10.260	10.010
Skewness	1.090	0.870	0.450
Kurtosis	4.840	5.020	4.520
JB	742	644	283
Probability	0.000	0.000	0.000
Correlation	$ln\bar{F}_{\mathrm{M}}$	$ln\bar{F}_{\mathrm{Q}}$	$ln\bar{F}_{\mathrm{Y}}$
$ln\bar{F}_{\mathrm{M}}$	1.00		
$ln\bar{F}_{\mathrm{Q}}$	0.94	1.00	
$ln\bar{F}_{\mathrm{Y}}$	0.77	0.91	1.00
Panel B: $\nabla ln\bar{F}$			
Mean	0.000	0.000	0.000
Median	0.000	0.000	0.000
Maximum	0.130	0.080	0.070
Minimum	-0.080	-0.070	-0.060
Std. Dev.	0.013	0.011	0.008
Skewness	0.330	0.090	0.330
Kurtosis	9.810	9.610	12.610
JB	4251	3973	8426
Probability	0.000	0.000	0.000
Correlation	$\nabla ln\bar{F}_{\mathrm{M}}$	$\nabla ln\bar{F}_{\mathrm{Q}}$	$\nabla ln\bar{F}_{\mathrm{Y}}$
$\nabla ln\bar{F}_{\mathrm{M}}$	1.00		
$\nabla ln\bar{F}_{\mathrm{Q}}$	0.84	1.00	
$\nabla ln\bar{F}_{\mathrm{Y}}$	0.68	0.89	1.00

Notes: This table reports some descriptive summary statistics for average swap prices computed as the geometric average of daily swap prices within each market segment and over the full sample period from 6/1/2004 to 12/31/2012. The sample size is 2179 observations. JB is the Jarque-Bera normality test. Probability is the p-value of this test under the null of normality.

3.4.2 Seasonal Components

We show the seasonal component computed using the method described in Equation (3.10) in Figure 3.3.

Figure 3.3: Average Swap Prices



Notes: The figure presents the estimated seasonal components $\widehat{s}_i(K) = \frac{1}{n}\sum_{T_k \in A_i}\sum_{t=1}^{n}\left(lnF_i(t,T_k) - ln\overline{F}_i(t)\right)$, where $A_{(i)}$ are the sets of available maturities at time t for contracts with delivery period $i = 1, \ldots, I$. And the index K refers to each month ($k = 1 \ldots 12$). Summary statistics for these seasonal components are below the plot. The last 3 columns present an interval of confidence for them and the t-test of the null hypothesis of zero mean.

As expected, swaps expiring in fall and winter are at a premium with respect to the average price level, and swaps expiring in spring and summer are at a discount. The January and February premium is the highest, at 9%. On the other hand, May has the highest discount, at 11%. A statistical significance test (not shown) reveals that most seasonal components are significantly different from zero at the usual levels.

3.4.3 Stochastic forward premium factors

Using Equation (3.11) we compute the estimated SFPs. Their basic statistics are in Table 3.3, in levels $\widehat{\gamma}_i(t,T)$ in Panel A and in first differences $\nabla\widehat{\gamma}_i(t,T)$ in Panel B. All $\widehat{\gamma}_i(t,T)$ series have means that are close to zero, as expected, and volatilities generally fall as we include more distant market segments, furthermore, all series present some asymmetry and kurtosis, as well as first order autocorrelation coefficients close to one (not shown). All series $\nabla\widehat{\gamma}_i(t,T)$ present means that are essentially zero, and volatilities that usually decrease with the time to maturity. Some present positive skewness but others present negative skewness, and the first-order autocorrelation coefficient (not shown) is usually below 0.1, suggesting a slow mean-reverting behaviour. High kurtosis seems to be a salient feature in all cases. Therefore, the evidence suggests that all these series are highly non-normal.

Table 3.3: Summary Statistics Stochastic Forward Premium (SFPs)

	M1	M2	M3	M4	M5	M6	Q1	Q2	Q3	Q4	Q5	Q6	Y1	Y2	Y3	Y4	Y5	Y6
Panel A: SFPs Levels $\gamma_i(T)$																		
Mean	-0.037	-0.014	-0.001	0.009	0.020	0.023	-0.036	-0.008	0.005	0.007	0.011	0.020	-0.048	-0.036	-0.019	0.013	0.037	0.053
Median	-0.033	-0.018	-0.003	0.008	0.018	0.021	-0.028	-0.004	0.017	-0.002	0.011	0.034	-0.036	-0.030	-0.018	0.005	0.030	0.049
Maximum	0.219	0.208	0.176	0.189	0.191	0.312	0.257	0.113	0.241	0.205	0.157	0.248	0.118	0.025	0.021	0.094	0.125	0.131
Minimum	-0.310	-0.215	-0.153	-0.101	-0.180	-0.251	-0.328	-0.178	-0.119	-0.086	-0.136	-0.187	-0.244	-0.110	-0.058	-0.037	-0.037	-0.032
Std. Dev.	0.083	0.060	0.044	0.043	0.060	0.084	0.098	0.068	0.044	0.047	0.067	0.076	0.065	0.029	0.013	0.028	0.034	0.037
Skewness	-0.389	-0.228	-0.067	0.326	0.161	0.031	-0.438	-0.321	-0.922	0.593	0.073	-0.242	-0.668	-0.234	-0.123	1.070	0.458	0.120
Kurtosis	3.532	3.570	3.795	3.574	3.366	3.690	3.028	2.265	3.616	3.362	2.093	2.836	3.101	2.350	3.244	3.486	2.572	2.281
JB	**80.69**	**48.31**	**58.98**	**68.46**	**21.57**	**43.55**	**69.74**	**86.56**	**343.20**	**139.62**	**76.62**	**23.74**	**162.96**	**58.19**	**10.89**	**437.51**	**92.65**	**52.19**
Panel B: First Differences of SFPs $\nabla\widehat{\gamma}_i(t,T)$																		
Mean	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Median	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	0.076	0.040	0.033	0.041	0.027	0.046	0.063	0.026	0.030	0.024	0.027	0.032	0.022	0.022	0.018	0.017	0.020	0.048
Minimum	-0.077	-0.064	-0.023	-0.051	-0.034	-0.041	-0.043	-0.024	-0.020	-0.024	-0.033	-0.036	-0.040	-0.023	-0.015	-0.016	-0.022	-0.032
Std. Dev.	0.011	0.006	0.005	0.006	0.007	0.008	0.007	0.004	0.004	0.004	0.005	0.005	0.005	0.003	0.003	0.003	0.003	0.004
Skewness	-0.064	-0.101	0.171	-0.243	-0.087	0.270	0.320	0.006	0.511	0.016	-0.067	0.111	-0.503	-0.150	0.576	0.161	-0.044	0.755
Kurtosis	9.812	12.36	6.156	8.610	5.160	6.650	11.546	7.480	8.961	7.007	7.344	6.521	7.942	8.493	9.466	6.747	10.36	22.07
JB	**80.69**	**48.31**	**58.98**	**68.46**	**21.57**	**43.55**	**69.74**	**86.56**	**343.20**	**139.62**	**76.62**	**23.74**	**162.96**	**58.19**	**10.89**	**437.51**	**92.65**	**52.19**

Notes: This table reports some descriptive summary statistics for SFPs, in levels $\gamma_i(t,T)$ (Panel A) and in first differences $\nabla\gamma_i(t,T)$ (Panel B) for the full sample period from 6/1/2004 to 12/31/2012. The sample size is 2179 observations. JB is the Jarque-Bera normality test. Boldface means that the test rejects the null of normality.

We show the correlations between the returns of the average series and the returns of the SFPs series in Table 3.4. The monthly contracts tend to be correlated with their average factor and with nearby monthly contracts as well as with some quarterly contracts. However, their correlation with the yearly contracts is usually not high. There is (relatively high) correlation across yearly contracts and there is correlation between the yearly contracts and its average factor but the correlation of the yearly contracts with the monthly and quarterly contracts is usually quite low. Overall, the evidence is consistent with the assumptions in our theoretical model, because it includes correlations both within and across market sectors, and thus it permits a more realistic representation of market prices.

Table 3.4: Correlations Average factors and SPFs

	M1	M2	M3	M4	M5	M6	Q1	Q2	Q3	Q4	Q5	Q6	Y1	Y2	Y3	Y4	Y5	Y6	$\nabla ln\bar{F}_{\mathrm{M}}$	$\nabla ln\bar{F}_{\mathrm{Q}}$	$\nabla ln\bar{F}_{\mathrm{Y}}$
M1	1.00																				
M2	**0.42**	1.00																			
M3	**-0.30**	**-0.09**	1.00																		
M4	**-0.56**	**-0.41**	**0.09**	1.00																	
M5	**-0.59**	**-0.52**	-0.05	**0.22**	1.00																
M6	**-0.60**	**-0.54**	**-0.14**	**0.16**	**0.29**	1.00															
Q1	**0.51**	**0.44**	**0.11**	**-0.31**	**-0.44**	**-0.52**	1.00														
Q2	0.00	-0.05	**-0.14**	-0.04	0.07	**0.10**	**0.19**	1.00													
Q3	**-0.20**	**-0.19**	0.01	**0.09**	**0.15**	**0.22**	**-0.37**	0.02	1.00												
Q4	**-0.18**	**-0.15**	0.01	**0.11**	**0.11**	**0.17**	**-0.48**	**-0.34**	0.10	1.00											
Q5	**-0.21**	**-0.14**	0.01	**0.14**	**0.18**	**0.14**	**-0.46**	**-0.45**	**-0.19**	**0.11**	1.00										
Q6	**-0.23**	**-0.20**	-0.07	**0.18**	**0.20**	**0.23**	**-0.50**	**-0.40**	**-0.12**	0.00	**0.14**	1.00									
Y1	**0.18**	**0.15**	0.07	**-0.10**	**-0.19**	**-0.19**	**0.15**	-0.01	0.02	0.05	-0.03	**-0.24**	1.00								
Y2	-0.01	0.03	0.02	0.01	-0.03	-0.01	-0.09	-0.06	0.05	0.05	**0.11**	-0.01	**0.57**	1.00							
Y3	-0.08	-0.06	-0.05	0.03	**0.11**	0.08	-0.10	0.00	0.01	-0.01	0.04	**0.11**	**-0.14**	**0.20**	1.00						
Y4	-0.08	-0.08	-0.04	0.05	0.08	**0.09**	-0.07	0.01	-0.01	0.00	0.00	0.10	**-0.49**	**-0.39**	0.05	1.00					
Y5	-0.05	-0.06	-0.03	0.01	0.06	**0.09**	0.00	0.05	-0.03	-0.06	-0.05	0.06	**-0.66**	**-0.70**	**-0.32**	**0.13**	1.00				
Y6	-0.06	-0.05	-0.01	0.05	0.07	0.04	0.00	0.00	-0.04	-0.05	-0.03	0.09	**-0.60**	**-0.68**	**-0.39**	-0.02	**0.58**	1.00			
$\nabla ln\bar{F}_{\mathrm{M}}$	**0.46**	**0.36**	-0.06	**-0.32**	**-0.34**	**-0.37**	**0.52**	**0.23**	**-0.13**	**-0.25**	**-0.29**	**-0.37**	**0.60**	**0.30**	**-0.14**	**-0.34**	**-0.35**	**-0.30**	1.00		
$\nabla ln\bar{F}_{\mathrm{Q}}$	**0.17**	**0.15**	0.02	**-0.11**	**-0.13**	**-0.18**	**0.17**	**0.11**	-0.01	-0.06	-0.08	**-0.21**	**0.69**	**0.44**	**-0.13**	**-0.42**	**-0.45**	**-0.38**	**0.84**	1.00	
$\nabla ln\bar{F}_{\mathrm{Y}}$	0.06	0.07	-0.02	-0.04	-0.02	-0.08	-0.02	0.01	0.02	0.04	0.04	-0.06	**0.40**	**0.29**	**-0.11**	**-0.30**	**-0.26**	**-0.18**	**0.69**	**0.89**	1.00

Notes: This table presents the correlations between the returns of the average factors $\nabla ln\overline{F}_i(t)$ and the returns of the SPF factors $\nabla\gamma_i(t,T)$ for each segment. The sample size is 2178 observations. Boldface indicates statistical significance at 5% level.

3.4.4 The relative weight of the components

An important practical question is the proportion of the total variation explained by each component for the different market segments. To study this issue we run a regression, where the explanatory variables are the components that are included sequentially. In the monthly segment, on average, the average swap price explains around 75% of the total variation, the SFP component explains an additional 15% and the seasonal component explains the remaining 10% of the total variation, and the differences across maturities are not very marked. The same situation appears in the quarterly segment, where the average swap price explains around 80%, the SFP component explains an additional 5% and the seasonal component explains the remaining 15% of the total variation. In the case of the yearly segment, and on average, the average swap price explains around 85%, and the SFP component explains the remaining 15% of the total variation.

One implication of these results is that the specific component in each swap contract represents a non-negligible source of risk. This risk is specific of each contract and we cannot hedge this risk by using other swap contracts. For instance if one trader wants to hedge a position in a yearly contract using monthly contracts, there is an specific risk that can be as high as more than 20% of total risk, and this idiosyncratic risk cannot be avoided by using monthly contracts. Additionally, by using these monthly contracts, the trader is assuming additional idiosyncratic risks in each monthly contract. The consequence of this is that a trader wishing to hedge some electricity swap contracts in the EEX market using other (shorter maturity) electricity swap contracts assumes significant basis risk.

3.4.5 Model Estimation

In Table 3.5 we present the results of the results of the estimation of model (3.12) for the average swap prices and the stochastic discount factors. Panel A contains the results for the average series factor and the estimation of their individual NIG distributions. Panel B contains the estimation of the parameters of the MNIG distribution of the standardized and orthogonalized residuals from Equation (3.12). Panels C and D contain the same information for the SFP of contract within each market segment (monthly, quarterly and yearly). The mean reversion parameter is significant in all cases, suggesting a mean reverting process, albeit with different speeds. The process is faster in the case of the SPF of the monthly contracts, followed by quarterly and then yearly contracts, being the slowest in the case of the processes for the average prices. The residual volatility varies considerably, being higher in the case of the residuals of the processes for the average prices and decreasing with the time to maturity in the case of the SPF across market segments, as expected.

A test of multivariate Normality (Urzua, 1997) of the residuals (not show) clearly reject the null of normality. Regarding the NIG parameters of the individual distributions, estimated α are always significant and vary from 0.44 (Y6) to 1.04 (Y2) , β are usually positive but non-significant, the δ are positive and significant, ranging from 0.43 to 1.03 (same contracts as before) , and the μ are very close to zero, as expected. The scale-free indicator of asymmetry, χ varies around 0.03 but usually it is not significant and the kurtosis parameter ξ is highly significant and varies from 0.68 (Y2) to 0.95 (Y6). The likelihood ratio test strongly supports the NIG distribution as a better alternative than the standard normal distribution for all individual distributions. It is also interesting to note that the SFP component corresponding to each swap contract, shares some common characteristics with the other SFPs within each market segment (symmetry, the degree of mean reversion) but it also has some specific features (different volatility levels).

The parameters of the MNIG distribution shown in Panel C give a similar message, with non-significant mean and asymmetry parameters and significant values in the cases of the parameters measuring tail heaviness and scale. The overall impression is that we can characterize the MNIG distribution as being essentially symmetric and having strong tail heaviness.

Table 3.5: Estimation of SURE and MNIG Models

	M	Q	Y
Panel A: Estimation $\nabla ln\overline{F}_i(t)$ and Marginal NIG			
κ_i	0.001**	0.001**	0.001**
ζ_i	4.062**	4.060**	4.0560**
$\theta_{\overline{F}_i}$	0.013	0.011	0.008
α	0.604**	0.538**	0.663**
β	0.030	-0.026	0.067*
δ	0.600**	0.503**	0.628**
μ	-0.030	0.025	-0.063
χ	0.045	-0.046	0.090*
ξ	0.906**	0.934**	0.886**
LR	41119**	37137**	29287**
Panel B: Parameters of MNIG residuals			
α	0.694**	0.694**	0.694**
β	0.046	0.035	0.028
δ	0.684**	0.684**	0.684**
μ	-0.017	-0.027	0.047

	M1	M2	M3	M4	M5	M6	Q1	Q2	Q3	Q4	Q5	Q6	Y1	Y2	Y3	Y4	Y5	Y6
Panel C: Estimation $\nabla\widehat{\gamma}_i(t,T)$ and Marginal NIG																		
$-\varpi_{i,T}$	-0.011**	-0.0142**	-0.033**	-0.032**	-0.017**	-0.011**	-0.0127**	-0.014**	-0.023**	-0.019**	-0.012**	-0.010**	-0.006**	-0.006**	-0.007**	-0.006**	-0.007**	-0.006**
$\theta_{\gamma_i(T)}$	0.011	0.006	0.005	0.006	0.006	0.008	0.0072	0.0041	0.004	0.004	0.005	0.005	0.005	0.003	0.002	0.003	0.003	0.004
α	0.711**	0.751**	0.846**	0.772**	0.763**	0.599**	0.7268**	0.8878**	0.917**	0.998**	0.916**	0.798**	0.828**	1.040**	0.997**	0.625**	0.498**	0.444**
β	0.038	0.041	0.039	-0.004	-0.001	0.068*	-0.0252	0.0339	0.016	0.048	0.071*	0.031	-0.048	0.007	0.080*	0.052*	0.018	0.027
δ	0.634**	0.655**	0.771**	0.679**	0.703**	0.569**	0.7002**	0.8953**	0.915**	0.997**	0.915**	0.810**	0.814**	1.031**	0.974**	0.640**	0.501**	0.438**
μ	-0.034	-0.035	-0.035	0.004	0.001	-0.064	0.0243	-0.0342	-0.016	-0.048	-0.071	-0.031	0.047	-0.007	-0.078	-0.053	-0.018	-0.027
χ	0.046	0.046	0.037	-0.005	-0.001	0.103*	-0.0296	0.0293	0.013	0.034	0.058*	0.031	-0.047	0.004	0.057*	0.074*	0.033	0.060*
ξ	0.871**	0.855**	0.803**	0.84**	0.842**	0.912**	0.854**	0.766**	0.752**	0.709**	0.753**	0.812**	0.802**	0.688**	0.714**	0.896**	0.943**	0.960**
LR	31519**	37140**	23287**	14753**	17695**	18847**	29519**	41137**	38767**	16113**	19655**	14567**	37779**	48127**	35541**	14432**	16755**	14517**
Panel D: Parameters of MNIG residuals																		
α	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**	0.694**
β	0.011	0.001	0.025	-0.025	-0.008	0.045	-0.013	0.009	0.029	0.032	0.033	0.039	-0.048	0.005	0.019	0.034	0.016	0.011
δ	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**	0.684**
μ	0.012	0.002	0.001	0.0345	0.008	0.005	0.011	0.01	0.0243	-0.034	-0.016	-0.048	-0.071	0.031	0.047	0.007	0.078	0.053

Notes: This table presents estimates of model (3.12). The SURE system contains 21 equations (3+63) and is estimated using the feasible generalized least squares (FGLS) method. To take into account residual autocorrelation a first order AR(1) term is included in all equations. In the second step we use the standardized residuals from Equation (3.12) $\epsilon_{\overline{F}_i}$ and $\epsilon_{\gamma_i(T)}$ using the Urzua (1997) method as estimations of the vector $dL(t)$, which is MNIG. The sample period spans from 6/1/2004 to 12/31/2012. The sample size is 2178 observations. Standard errors for the parameters in the first step are heteroskedasticity- and autocorrelation consistent. Standard errors for the NIG parameters are based on generating 500 bootstrapped samples. * denotes significance at 5% level and ** at 1% level. The likelihood ratio test is computed as follows $LR = -2 * LOGLK(N(0,1)) + 2 * LOGLN(NIG_d(\alpha, \beta, \delta, \mu))$. For each equation and marginal distribution. Panel A contains the results for the average series factor and the estimation of their marginal NIG distributions. Panel B contains the estimation of the parameters of the MNIG distribution of the standardized and orthogonalized residuals from Equation (3.12). Panels C and D contain the same information for the SFP of contract within each market segment (monthly, quarterly and yearly).

Table 3.6 contains the correlations among the ordinary residuals from Equation (3.12) and in boldface; we highlight the correlations used in the computations of the volatility term structure.

Table 3.6: Correlations of the residuals of model (12)

	M1	M2	M3	M4	M5	M6	Q1	Q2	Q3	Q4	Q5	Q6	Y1	Y2	Y3	Y4	Y5	Y6	$\nabla ln\bar{F}_{M}$	$\nabla ln\bar{F}_{Q}$	$\nabla ln\bar{F}_{Y}$
M1	1.00																				
M2	0.42	1.00																			
M3	-0.30	-0.09	1.00																		
M4	-0.57	-0.42	0.10	1.00																	
M5	-0.60	-0.52	-0.05	0.23	1.00																
M6	-0.60	-0.54	-0.14	0.16	0.30	1.00															
Q1	0.51	0.44	0.11	-0.31	-0.44	-0.53	1.00														
Q2	0.00	-0.04	-0.14	-0.05	0.06	0.10	0.19	1.00													
Q3	-0.21	-0.19	0.00	0.10	0.16	0.24	-0.37	0.02	1.00												
Q4	-0.19	-0.15	0.01	0.12	0.12	0.18	-0.49	-0.35	0.10	1.00											
Q5	-0.22	-0.14	0.01	0.15	0.18	0.15	-0.47	-0.46	-0.18	0.12	1.00										
Q6	-0.24	-0.21	-0.07	0.18	0.21	0.24	-0.52	-0.41	-0.11	0.01	0.16	1.00									
Y1	0.18	0.15	0.08	-0.10	-0.19	-0.20	0.16	0.00	0.01	0.04	-0.03	-0.25	1.00								
Y2	-0.01	0.02	0.02	0.01	-0.01	-0.01	-0.09	-0.06	0.05	0.05	0.11	0.00	0.56	1.00							
Y3	-0.08	-0.06	-0.05	0.03	0.11	0.08	-0.11	-0.01	0.01	0.00	0.04	0.12	-0.15	0.21	1.00						
Y4	-0.08	-0.08	-0.04	0.05	0.08	0.10	-0.07	0.00	0.00	0.00	0.00	0.11	-0.51	-0.40	0.07	1.00					
Y5	-0.05	-0.06	-0.03	0.02	0.05	0.09	0.00	0.05	-0.02	-0.06	-0.05	0.06	-0.67	-0.71	-0.32	0.16	1.00				
Y6	-0.06	-0.05	-0.02	0.05	0.07	0.04	0.00	0.01	-0.03	-0.05	-0.03	0.09	-0.61	-0.68	-0.39	0.00	0.59	1.00			
$\nabla ln\bar{F}_{M}$	**0.47**	**0.37**	**-0.06**	**-0.32**	**-0.35**	**-0.38**	0.53	0.23	-0.14	-0.25	-0.29	-0.38	0.60	0.31	-0.14	-0.35	-0.35	-0.31	1.00		
$\nabla ln\bar{F}_{Q}$	0.17	0.15	0.02	-0.11	-0.14	-0.18	**0.18**	**0.11**	**-0.01**	**-0.07**	**-0.08**	**-0.22**	0.69	0.45	-0.13	-0.44	-0.46	-0.38	0.84	1.00	
$\nabla ln\bar{F}_{Y}$	0.06	0.07	-0.02	-0.04	-0.02	-0.08	-0.02	0.01	0.02	0.04	0.04	-0.06	**0.40**	**0.30**	**-0.11**	**-0.31**	**-0.27**	**-0.19**	0.69	0.89	1.00

Notes: This table presents the correlations of the residuals from (3.12). The sample size is 2178 observations. Boldface indicates the correlations needed in order to compute the term structure of swap prices given by (3.8).

3.4.6 In-sample Goodness-of-Fit

As a formal test of the extent to which the NIG distribution is successful in representing the (one-day) innovations in the marginal distributions of the average swap price and the SFPs, we implement tests of fit based on the empirical distribution function (EDF). These statistics measure the discrepancy between the EDF and a given theoretical distribution (e.g. Normal or NIG). We calculate two statistics: the Cramer-von Mises and the Kolmogorov-Smirnov. We calculate the parameters of the NIG distribution by maximum likelihood. Next, we face the problem of how to evaluate these test statistics, given that the true parameter values of the (NIG) distribution are unknown. To solve this problem, and following Capasso, Alessi, Barigozzi, and Fagiolo (2009), we generate 5000 Monte Carlo simulations of i.i.d. NIG random numbers for each market segment. In doing so, we obtain approximate distributions of the EDF test statistics. We briefly summarize the results. The null hypothesis is that the stochastic elements in the marginal distributions follow NIG distributions. The number of violations of the null hypothesis at the 5% significance level is always lower than the critical value suggesting that, in all cases, one-day returns follow NIG distributions.

3.4.7 Volatility term structure

In Table 3.7 we present the volatility term structure that has been computed by using Equation (3.8) and (3.12) with the calibrated parameters obtained in Table 3.5 and Table 3.6.

Table 3.7: Correlations of the residuals of model (12)

	Market	Model	Absolute Error	Relative Error
M1	0.3264	0.3269	-0.30%	0.30%
M2	0.2613	0.2620	-0.59%	0.59%
M3	0.2190	0.2190	-0.03%	0.03%
M4	0.2002	0.2003	-0.17%	0.17%
M5	0.1999	0.1991	0.76%	0.76%
M6	0.1997	0.1988	0.94%	0.94%
Q1	0.2245	0.2248	-0.32%	0.32%
Q2	0.1925	0.1928	-0.29%	0.29%
Q3	0.1831	0.1830	0.12%	0.12%
Q4	0.1815	0.1813	0.26%	0.26%
Q5	0.1843	0.1840	0.28%	0.28%
Q6	0.1759	0.1753	0.68%	0.68%
Y1	0.1738	0.1739	-0.16%	0.16%
Y2	0.1506	0.1511	-0.72%	0.72%
Y3	0.1311	0.1312	-0.12%	0.12%
Y4	0.1237	0.1231	1.06%	1.06%
Y5	0.1264	0.1259	0.90%	0.90%
Y6	0.1336	0.1330	0.83%	0.83%
Average			0.17%	0.47%

Notes: This table presents the market volatility of swap returns and estimated volatility using the term structure of swap prices variances. The term structure of swap prices variances is given by $\varphi_i^2(t,T) = \eta_i^2 + \tau_{i,T}^2 + 2\theta_{\overline{F}_i}\theta_{\gamma_i(T)}\upsilon_{\overline{F}_i,\gamma_i(T)}$.

As may be seen in Table 3.7 the model is able to reproduce the overall volatility structure with high degree of precision irrespective of the market segment. Average absolute and relative errors are lower than 0.5%. The average contributions of each of the elements in (3.8) to the total variance explained by the model are as follows: 87.5% corresponds to the variance of the corresponding average factor, 18.5% corresponds to the variance of the stochastic discount factor, and -6.1% corresponds to the covariance term between the average factor and the stochastic discount factor. However the range of variation is substantial, being in the first case from 41% to 111%, in the second case from 9% to 37% and in the third case from -49% to 31%. The implication of these results is that the order of magnitude of determinants of the variance for each contract should be analysed carefully because, although in some cases the impact of the average factor dominates, in other cases the impact of the covariance component can be determinant.

3.5 Comparison against alternative models

Here we compare the performance of the baseline model introduced in Section 3.2 against two competing models, the first one is a one-factor spot price model and the second one is a HJM-based approach. Details are in Appendix B.1 and Appendix B.2 respectively.

3.5.1 Spot Prices: One factor Model

The results of the estimation of this model and the pricing exercise with the computation of RMSE are in Appendix B.1. As an illustration, we present the results for contracts M1, Q1 and Y1 (which are the liquid contracts within each market segment) during 2010 and compare them against market prices. The results suggest that this model is not able to capture the basic features of swap market prices. In particular, theoretical prices are much more volatile than market prices as the graphs in Figure B.2 suggest. The reason is that theoretical prices are simple functions of spot prices which are always much more volatile than swap prices. The correlation between theoretical prices and market prices is not particularly high, being (on average) 0.53, 0.41 and 0.09, for monthly, quarterly and yearly swaps respectively. The pricing errors are not independent and the theoretically based variances do not appropriately reflect market volatility. Besides that, the residuals from the model are strongly non-normal, contradicting the basic assumptions underlying this model. In summary, this one-factor model does not capture the basic characteristics of the swap prices and therefore is unlikely to be useful for pricing or hedging purposes.

3.5.2 Swap Prices: HJM Model

The results of the estimation of this model are in Appendix B.2. The eigenvalues resulting from the eigenvector decomposition tell us the importance of each eigenvector and hence the number of factors that we should include in our model. Thus, the first eigenvector is the most important, explaining 87%, 89% and 94% of the total variation in the evolution of the swap curve for the monthly, quarterly and yearly contracts respectively.

Table B.3 presents in Panel A the LS estimates of the parameters from the volatility functions obtained in CPA using equation $\sigma_{1i}(t, T_i) = \exp -k_i(T_i - t)\sigma_{1i}$ for the full sample period 2004-2012. Panel B reports the in-sample root mean squared pricing errors (RMSEs). We compute daily errors based on the fitted swap prices in Panel A. We compute the volatility function implied by the HJM model and by the SFP model and compare the results against market prices. In the case of the HJM model, root-mean squared errors (RMSEs) are 6.05%, 19.32%, and 11.96% for monthly, quarterly and yearly contracts respectively. By contrast, the RMSEs for those contracts are 0.10%, 0.12% and 0.30% respectively in the case of the SFP model. The degree of fit of the SFP model is substantially higher than the HJM model. It is fair to say that, although the model seems to fit the volatility term structure to some extent, it is unable to recover the skewness and kurtosis observed in the empirical distributions. By contrast, the SFP model not only fits better the markets volatility term structure, but also is able to take into account the skew and kurtosis.

3.6 Value at Risk

In order to compare the performance of alternative models in comparison with the SFP model, we compute Value-at-Risk (VaR) at different probability levels over a one-day horizon. The details of the procedure are in Appendix B.3. We present the results in Table 3.9. It may be seen that the VaR_Q^{Normal} , calculated under the assumption of normality tends to understate risk, and this understatement is very strong for high confidence levels (99.5% and 99.99%) suggesting that tail risk is severely underestimated. On the other hand VaR_Q^{NIG} tends to mildly overestimate risk at relatively low significance levels but it is able to properly account for extreme tail risk. It is worth noting that the overestimation of risk provided by the NIG distribution is proportionally much lower than the underestimation of risk produced by the normal distribution.

One important practical implication of our results is as follows. We know that VaR models allow users to control risk and decide how to allocate limited resources. Financial intermediaries impose a capital charge to traders based on risk-adjusted capital. This creates a natural incentive for traders to take a position only when they have strong views on markets. If they have no views, they should abstain from trading. Our results suggest that the risk-adjusted capital for traders using EEX swap electricity contracts should be increased in comparison with the standard practice based on the normality assumption. Traders should also rationally adjust positions as risk changes (in the face of an increasingly volatile environment a sensible response is to scale down positions). Furthermore and given that VaR is also a performance evaluation tool, the evaluators of the performance of the traders should adjust their measures accordingly.

Table 3.8: Correlations of the residuals of model (12)

	N(1%)	N(0.5%)	N(0.01%)	NIG(1%)	NIG(0.5%)	NIG(0.01%)
M1	2.84**	3.07**	14.53**	-1.24	-1.79	1.68
M2	3.27**	4.59**	8.10**	-2.32**	-2.39**	1.68
M3	2.84**	5.50**	10.24**	-3.18**	-3.30**	-0.47
M4	2.63**	3.98**	14.53**	-3.39**	-3.30**	-0.47
M5	3.70**	4.89**	16.67**	-3.18**	-3.30**	-0.47
M6	4.35**	6.43**	18.81**	-3.18**	-3.30**	-0.47
Q1	3.70**	6.10**	10.24**	-3.18**	-2.70**	-0.47
Q2	4.56**	6.41**	18.81**	-4.04**	-3.30**	-0.47
Q3	3.70**	6.41**	16.67**	-2.96**	-2.39**	-0.47
Q4	2.41**	3.67**	18.81**	-2.96**	-3.00**	-0.47
Q5	4.35**	6.41**	23.10**	-3.83**	-2.70**	-0.47
Q6	5.21**	7.93**	14.53**	-3.83**	-2.70**	1.68
Y1	4.35**	6.71**	16.67**	-0.17	-1.49	1.68
Y2	4.13**	6.71**	10.24**	-1.46	-2.39**	1.68
Y3	3.70**	6.10**	10.24**	-1.68	-2.70**	1.68
Y4	2.63**	3.67**	10.24**	-3.18**	-2.70**	1.68
Y5	2.63**	5.50**	20.96**	-1.89	-2.39**	-0.47
Y6	3.70**	6.10**	25.24**	-1.68	-2.09*	-0.47

Notes: This table reports the results of Wald test for daily swap prices. The sample period spans from 6/1/2004 to 12/31/2012. The sample size is 2178 observations. **, or * indicates that the coefficient estimate is significantly different from the null hypotheses at the 1%, or 5% level, respectively. A positive (negative) z statistic indicates that the model underestimates (overestimates) risk.

3.7 Conclusions

It is important to develop models of electricity markets that are able to handle the non-Gaussian nature of electricity prices. The reason is that risk measures and derivatives prices depend on the distribution used in a model. From a structural point of view, we posit that swap electricity prices result from at least three driving forces. First, a stochastic factor acting as an anchor of the overall level of the curve, this factor representing the average consensus price for the contracts within a given maturity slot (yearly, quarterly, monthly). Second, a factor reflecting stable seasonal components because of market anticipation of weather-related variations in demand. Third, a factor accounting from (men-reverting) stochastic deviations from the previous two factors, these deviation depending on the time to maturity and the length of the delivery period.

Besides that, the innovations of the stochastic factors are non-normal, a critical fact that should be taken properly into account. In particular, failure for account for asymmetries and fat tails leads to theoretical prices that are not compatible with market prices. Our model takes into account all these features. With an empirical application based the largest European power market (Germany), during the period from 2004 to 2013, the evidence strongly support our model.

By accounting for all these factors, our model gains extra explanatory power, compared with a one-factor model based on spot prices and with a HJM-based model. Another remarkable point is the ability our model has in capturing extreme tail risk, as suggested by the results of the VaR analysis. A practical implication of our results is that the capital charges to traders using EEX electricity contracts, based on risk-adjusted capital and on the normality assumption, are too conservative. We suggest increasing them and that the evaluators of the performance of the traders should adjust their recommendations accordingly.

Looking forward, an application of our model to other electricity markets offers an interesting topic for further investigation, given the local behaviour of many electricity markets. The application of our model for the pricing of structured derivative contracts, hedging strategies, portfolio diversification, and risk management purposes represent other natural directions for further research.

Appendix A

Appendix for "The Bright Side of Financial Derivatives: Options Trading and Firm Innovation"

This Appendix provides additional material to the results presented in *"The Bright Side of Financial Derivatives: Options Trading and Firm Innovation"*. In Section A.1, we describe the construction of the main dataset. In Section A.2, we discuss and report robustness tests for the baseline results reported in Section 1.4 of the paper. In Section A.3, we report additional tests which supplement other parts of the main article. Descriptive statistics are in Table A.16.

A.1 Main dataset

The main firm level data sample is generated through the combination of several datasets. Because we are using patents (weighted by total future citations) as our key measure of innovation, we rely on the matching of the United States Patent and Trademark Office (USPTO) to the North American Compustat data hosted at the National Bureau of Economic Research (NBER) (see Hall, Jaffe, and Trajtenberg, 2001; Jaffe and Trajtenberg, 2002, for details). The main matching was performed based on the concordance file provided by Bessen (2009) that connects the assignee identification number of the NBER patent data set to the Compustat GVKEY identification number. These connections reflected the firms and subsidiaries identified in the *Who Own's Whom?* database (published annually by Dun & Bradstreet International). Ownership may change through mergers, acquisitions, or spin-offs, and when an organization is acquired/merged/spun-off, its patents likely go to the new owner. These changes have been tracked using data on the mergers and acquisitions of public companies reported in the SDC database. We use the updated version of the NBER match containing citations through to 2006 (downloaded from the NBER Patent Data Project website, `https://sites.google.com/site/patentdataproject/Home/downloads`). All patents granted between 1976 and 2004 are included (just under 3 million patents) and citation information is available from 1976 to 2006 (over 23 million citations). The need to have some patent data is the main reason why our sample is much smaller than the full Compustat sample.

The second dataset we draw on comes from OptionMetrics LLC, a financial research firm specializing in the analysis of option markets. The IvyDB US data set from OptionMetrics contains daily closing option prices (bid and ask) for all U.S. exchange-listed and NASDAQ equities and market indices, as well as all U.S. listed index and equity options, starting from January 1996 (which is why this is the first year in our sample). Besides option prices, it also contains daily time-series of the underlying spot prices, dividend payments and projections, stock splits, historical daily interest rate curves and, most importantly, option volumes. Implied volatilities and sensitivities (delta, gamma, vega and theta) for each option are calculated as well. The comprehensive nature of the database makes it most suitable for empirical work on option markets. The primary key (Security ID) for all data contained in IvyDB is linked to the security's CUSIP number and ticker symbol, so the merging of the two datasets is straightforward.

Third, we obtain data on institutional ownership from Thomson Reuters' CDA/Spectrum Institutional Holdings dataset. Starting in 1978, all institutions with more than $100 million of securities under discretionary management are required to report their holdings to the Securities and Exchange Commission (SEC) using form 13F. Each quarter, these institutions must disclose any common stock positions greater than 10,000 shares or greater than $200,000 in value. The data includes the numbers of institutional owners, the number of share issues and the percentage of outstanding shares held by each institution. For each fiscal year, we take the average of the four quarterly institutional holdings given by form 13F and treat that as our measure of institutional ownership (*InstOwn*). Given that the ownership data does not cover all the firms in the dataset, we lose 304 firms when we match the Compustat accounting data and ownership data.

We started with the NBER USPTO/Compustat match and kept all domestic firms trading on NYSE (stock exchange code 11), AMEX (12) and NASDAQ (14) with non-missing accounting data on fixed assets (PPENT), employees (EMP) and sales (SALE) that are listed on Compustat for at least three years. Since our preferred regressions use fixed effects, we condition our sample on firms who had received at least one citation and had at least two years of non-missing data on all variables. This leaves us with a merged dataset of 1,329 firms and 9,265 observations between 1996 (the first year of the options data) and 2004 (the last year of the patent data). For reasons explained in the main article, our final sample consists of firms with positive options volume that are active in five broadly defined R&D-intensive industries: (i) pharmaceuticals (SIC code 283), (ii) industrial and commercial machinery and computer equipment (35), (iii) electronics and communications (36), (iv) transportation equipment (37), and (v) instruments and related products (38). This leaves us with 3,271 observations on 548 firms which is our baseline sample.

A.2 Robustness tests for the baseline results

We conduct a rich set of robustness tests for our baseline results and report them in Tables A.1 – A.11. First, we check whether our results are robust to alternative econometric models. We start with a Poisson model when the dependent variable is the number of cite-weighted patents and the number of (unweighed) patents and report the results in Table A.1. As shown, the coefficients on *Ln(Optvol)* remain positive and significant across all columns, consistent with our baseline findings. For example, the coefficient estimate on *Ln(Optvol)* is 0.143 (p-value < 0.01) if we reproduce our baseline fixed effects model of cite-weighted patents (column 4 of Table 1.2 in the main article) and is 0.106 (p-value < 0.05) when we use simple patent counts as dependent variable. Next, because our dependent variables are right skewed (e.g., 24% of our sample firms have a zero number of citations), we use three modelling strategies that take this into account. We report the results in Table A.2. In columns 1 and 2, we adopt a quantile regression approach at the 75$^{\text{th}}$ percentile. The baseline results continue to hold and we obtain similar findings if we run the quantile regressions at the 70$^{\text{th}}$, 80$^{\text{th}}$, 85$^{\text{th}}$ and the 95$^{\text{th}}$ percentiles. We then use zero-inflated negative binomial (columns 3 and 4) and zero-inflated Poisson models (columns 5 and 6). We also find consistent results.

Second, because our main analysis uses contemporaneous independent variables, we run alternative specifications where we lag the variables. As a first step, our approach has been to explore empirically the effects of time lags between options trading and the dependent variables. Estimating models with various time lags (i.e., from $t-1$ to $t-5$) for the options trading variable, we found broadly consistent results for all models, but with coefficients on *Ln(Optvol)* that were consistently larger compared to the ones obtained from the contemporaneous models. We present the results of models with one- and three-year lagged explanatory variables, since adding further lags reduce the number of observations for firms in the dataset, without providing any appreciable gain in the precision of the estimates. The coefficients on *Ln(Optvol)* are shown in panels A (one-year lag) and B (three-year lag) of Table A.3, and are positive in all regressions. For example, the coefficients in column 1 suggest that increasing options trading activity from the sample median ($8.5 million) to the 75$^{\text{th}}$ percentile ($53.5 million) is associated with a 98% increase in future cite-weighted patents in the following year and a 67% increase in three years, all significant at the 1% level.

Third, we examine if the effect of options volume on innovation is monotonic (i.e., after conditioning on covariates). In Table A.4, we begin with the inclusion of *Ln(Optvol)* and its squared term. We find that the impact of *Ln(Optvol)* on cite-weighted patents is still positive and significant (coefficient = 0.105 and p-value < 0.1 in column 1 of Table A.4), but the coefficient estimate on the squared term, *Ln(Optvol) x Ln(Optvol)*, is not significant. Next, we create a dummy variable, *High Optvol*, that equals one if the options volume for a given firm is above the median in that year and zero otherwise, and interact this dummy with *Ln(Optvol)*. We then re-estimate Eq. (1) in the main article by adding the *High Optvol* dummy and the interaction term, *Ln(Optvol) x High Optvol*. However, as shown in columns 2, 4 and 6, the coefficient estimates on the interaction terms are not statistically significant, while the coefficients on *Ln(Optvol)* remain positive and highly significant. In un-tabulated analyses, we obtain similar results if we replace the dependent variable with unweighed patent counts. Overall, and consistent with the bivariate relationship in Figure 1.1 in the main article, it appears that the effect of options trading activity on innovation is monotonic.

Fourth, our preferred control for R&D inputs is a continuous measure of the depreciated sum of past R&D expenditures. Although widely used in prior studies, it may partly conceal some of the effects of R&D. To mitigate such concerns, we include just the contemporaneous R&D flow and establish dummy variables based on deciles of the distribution of *R&D stock*. We report the results in Tables A.5 and A.6, respectively. In both cases, the coefficients on options volume continue to be positive and significant. For example, according to columns 1 and 2 of Table A.5, an increase in options volume from the sample median ($8.5 million) to the 75$^{\text{th}}$ percentile ($53.5 million) is associated with a 74% increase in citations and a 74% increase in the number of patents filed.

Fifth, given that our sample period (1996 – 2004) includes the "dot-com bubble" which is conventionally dated between 1996 and 2000 (Ljungqvist and Wilhelm, 2003), we rerun our regressions for the two sub-periods 1996 – 2000 and 2001 – 2004. As can be seen in Table A.7, the coefficients on *Ln(Optvol)* are positive and significant (at the 1% level) in both sub-periods, which provides reassurance that our results are not driven by high coefficient magnitudes in the earlier or later period.

Finally, in Table A.8 we report the regression results after the inclusion of additional (financial) control variables. In Table A.9 we report the regression results after controlling for firms' external knowledge acquisition activities. In Table A.10 we report the regression results on the differential effect of options trading on innovation in R&D- and non-R&D-intensive industries based on a matched sample. In Table A.11 we report the within-firm regression results that compare changes in innovation before and after firms' inclusion in options markets. These findings are discussed in Section 1.4.2 of the main article.

A.3 Other tests

In this section, we show additional regression results that supplement other parts of the paper. We discuss these results in the main text.

In Table A.12 we show the results from the two-stage least-squares regression (2SLS) using the average open interest across all options on a stock throughout the calendar year as alternative instrument.

In Table A.13 we show the regression results on the interaction between options volume and managerial entrenchment using the "Entrenchment Index" (E-Index) (see Bebchuk, Cohen, and Ferrell, 2009, for details).

In Table A.14 we show the regression results from examining the effect of firm innovation (and options volume) on its market valuation (*Tobin's Q*).

In Table A.15 we show the regression results from examining the effect of options volume on innovation after controlling for all five economic mechanisms.

Table A.1: Innovation and options volume – Poisson model

Dependent Var.	CITES				PATS			
Method: Poisson	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Optvol)	0.230***	0.139***	0.238***	0.143***	0.132***	0.117***	0.122**	0.106**
	(0.062)	(0.034)	(0.067)	(0.035)	(0.051)	(0.042)	(0.053)	(0.044)
InstOwn	-0.040	-0.088	-0.055	-0.102	-0.189	-0.083	-0.090	-0.054
	(0.240)	(0.221)	(0.219)	(0.215)	(0.237)	(0.235)	(0.223)	(0.222)
Ln(K/L)	0.634***	0.519***	0.676***	0.555***	0.477**	0.371**	0.530***	0.427***
	(0.232)	(0.164)	(0.252)	(0.170)	(0.186)	(0.165)	(0.195)	(0.165)
Ln(Sales)	0.531***	0.250***	0.219*	0.128*	0.610***	0.330***	0.214**	0.148**
	(0.091)	(0.067)	(0.117)	(0.076)	(0.070)	(0.065)	(0.095)	(0.069)
Ln(Age)	-0.042	-0.261**	-0.175*	-0.330***	-0.014	-0.255**	-0.191**	-0.352***
	(0.110)	(0.112)	(0.102)	(0.099)	(0.086)	(0.109)	(0.086)	(0.093)
Ln(R&D stock)			0.349**	0.165*			0.469***	0.275***
			(0.138)	(0.090)			(0.109)	(0.091)
Firm fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3271	3271	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*) and other firm-level control variables. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies and time dummies. The time period is 1996 – 2004 (with citations up to 2006); fixed effects are based on including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Innovation and options volume – Other (alternative) specifications

Method	Quantile regression		Zero-inflated NB		Zero-inflated Poisson	
Dependent Var.	Ln(1+CITES)	Ln(1+PATS)	CITES	PATS	CITES	PATS
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.105***	0.106***	0.164***	0.142***	0.155***	0.115**
	(0.016)	(0.013)	(0.028)	(0.024)	(0.040)	(0.049)
InstOwn	-0.034	-0.153*	0.046	0.102	-0.088	0.004
	(0.128)	(0.088)	(0.164)	(0.175)	(0.211)	(0.218)
Ln(K/L)	0.021	0.032	0.167**	0.182***	0.564***	0.449***
	(0.052)	(0.036)	(0.070)	(0.060)	(0.168)	(0.163)
Ln(Sales)	0.123***	0.118***	0.108***	0.138***	0.114	0.137**
	(0.025)	(0.021)	(0.038)	(0.043)	(0.074)	(0.068)
Ln(Age)	-0.167***	-0.082*	-0.215**	-0.235***	-0.318***	-0.356***
	(0.060)	(0.044)	(0.087)	(0.076)	(0.097)	(0.092)
Ln(R&D stock)	0.315***	0.262***	0.285***	0.266***	0.163*	0.276***
	(0.030)	(0.025)	(0.039)	(0.044)	(0.088)	(0.089)
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of quantile (at the 75^{th} percentile), zero-inflated NB and zero-inflated Poisson panel regressions of firms' patents weighted by the number of forward citations ($CITES$) and firms' unweighted patent counts ($PATS$) on options volume ($Optvol$) and other firm-level control variables. Firms in all columns: 548. Robust standard errors in columns 1 and 2 are obtained from 200 bootstrap replications. All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Innovation and options volume – Lagged explanatory variables

Method	OLS		NB		Poisson	
Dependent Var.	Ln(1+CITES)	Ln(1+PATS)	CITES	PATS	CITES	PATS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: One-year lag						
Ln(Optvol)	0.185***	0.179***	0.181***	0.152***	0.146***	0.123**
	(0.033)	(0.028)	(0.033)	(0.027)	(0.047)	(0.056)
InstOwn	-0.108	-0.215	0.002	0.025	-0.103	-0.024
	(0.182)	(0.162)	(0.202)	(0.172)	(0.215)	(0.204)
Ln(K/L)	0.023	0.042	0.078	0.151**	0.610***	0.463***
	(0.060)	(0.050)	(0.079)	(0.067)	(0.191)	(0.170)
Ln(Sales)	0.123***	0.126***	0.132***	0.152***	0.147	0.170**
	(0.044)	(0.036)	(0.043)	(0.040)	(0.090)	(0.072)
Ln(Age)	-0.116	-0.059	-0.215**	-0.258***	-0.353***	-0.385***
	(0.090)	(0.077)	(0.100)	(0.082)	(0.104)	(0.091)
Ln(R&D stock)	0.255***	0.200***	0.311***	0.265***	0.146	0.242**
	(0.050)	(0.046)	(0.047)	(0.045)	(0.105)	(0.101)
Observations	2658	2658	2658	2658	2658	2658
Panel B: Three-year lag						
Ln(Optvol)	0.130***	0.179***	0.138***	0.164***	0.157***	0.132**
	(0.042)	(0.039)	(0.044)	(0.033)	(0.049)	(0.061)
InstOwn	-0.058	-0.194	0.142	0.027	-0.064	0.099
	(0.222)	(0.201)	(0.243)	(0.192)	(0.265)	(0.213)
Ln(K/L)	0.005	0.048	0.179**	0.143**	0.733***	0.578***
	(0.068)	(0.062)	(0.084)	(0.070)	(0.222)	(0.174)
Ln(Sales)	0.159***	0.158***	0.186***	0.167***	0.153	0.205***
	(0.047)	(0.042)	(0.054)	(0.042)	(0.115)	(0.076)
Ln(Age)	-0.099	-0.105	-0.211*	-0.294***	-0.409***	-0.452***
	(0.099)	(0.091)	(0.112)	(0.083)	(0.120)	(0.092)
Ln(R&D stock)	0.215***	0.170***	0.259***	0.236***	0.118	0.182*
	(0.051)	(0.051)	(0.053)	(0.048)	(0.122)	(0.105)
Observations	1687	1687	1687	1687	1687	1687

Notes: This table presents estimates of OLS, NB and Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on (lagged) options volume (*Optvol*) and other (lagged) firm-level control variables. Firms in all columns: 526 in panel A and 399 in panel B. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Innovation and options volume – Monotonic relationship?

Method	OLS		NB		Poisson	
Dependent Var.	Ln(1+CITES)	Ln(1+CITES)	CITES	CITES	CITES	CITES
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)		0.065		-0.012		-0.045
x High Optvol		(0.069)		(0.067)		(0.072)
High Optvol		-0.159		0.042		0.166
		(0.269)		(0.251)		(0.259)
Ln(Optvol)	0.009		-0.003		-0.011	
x Ln(Optvol)	(0.008)		(0.008)		(0.008)	
Ln(Optvol)	0.105*	0.127***	0.176***	0.162***	0.254***	0.184***
	(0.059)	(0.044)	(0.060)	(0.046)	(0.074)	(0.059)
InstOwn	-0.017	-0.024	0.066	0.068	-0.126	-0.103
	(0.159)	(0.159)	(0.178)	(0.178)	(0.222)	(0.222)
Ln(K/L)	0.025	0.024	0.107	0.106	0.540***	0.553***
	(0.052)	(0.052)	(0.067)	(0.067)	(0.172)	(0.172)
Ln(Sales)	0.127***	0.127***	0.135***	0.135***	0.140*	0.128*
	(0.041)	(0.041)	(0.040)	(0.040)	(0.072)	(0.075)
Ln(Age)	-0.107	-0.106	-0.213**	-0.213**	-0.326***	-0.328***
	(0.084)	(0.084)	(0.095)	(0.095)	(0.102)	(0.102)
Ln(R&D stock)	0.253***	0.253***	0.301***	0.300***	0.159*	0.164*
	(0.046)	(0.046)	(0.042)	(0.042)	(0.085)	(0.086)
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS, NB and Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) on options volume (*Optvol*), its squared term, a dummy variable for high options volume (*High Optvol*), its interaction with options volume and other firm-level control variables. *High Optvol* equals one if the options volume for a given firm is above the median in year t and zero otherwise. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Innovation and options volume – Contemporaneous R&D spending

Method	OLS		NB		Poisson	
Dependent Var.	Ln(1+CITES)	Ln(1+PATS)	CITES	PATS	CITES	PATS
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.140***	0.136***	0.133***	0.122***	0.141***	0.091*
	(0.030)	(0.025)	(0.029)	(0.025)	(0.043)	(0.052)
InstOwn	-0.041	-0.218	0.021	-0.032	-0.084	-0.039
	(0.156)	(0.137)	(0.177)	(0.147)	(0.207)	(0.214)
Ln(K/L)	0.026	0.042	0.099	0.141**	0.562***	0.430***
	(0.052)	(0.044)	(0.065)	(0.058)	(0.171)	(0.164)
Ln(Sales)	0.100**	0.095***	0.131***	0.132***	0.078	0.094
	(0.042)	(0.034)	(0.042)	(0.039)	(0.070)	(0.064)
Ln(Age)	-0.040	0.024	-0.144	-0.151**	-0.298***	-0.291***
	(0.082)	(0.067)	(0.093)	(0.075)	(0.109)	(0.103)
Ln(1+XRD)	0.316***	0.256***	0.331***	0.290***	0.224**	0.349***
	(0.052)	(0.046)	(0.048)	(0.047)	(0.093)	(0.084)
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS, NB and Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*), contemporaneous R&D spendings (*XRD*) and other firm-level control variables. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Innovation and options volume – R&D stock dummy variables

Method	OLS		NB		Poisson	
Dependent Var.	Ln(1+CITES)	Ln(1+PATS)	CITES	PATS	CITES	PATS
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.163***	0.145***	0.149***	0.132***	0.147***	0.115***
	(0.029)	(0.024)	(0.028)	(0.024)	(0.032)	(0.043)
InstOwn	-0.088	-0.188	0.073	-0.039	-0.271	-0.187
	(0.166)	(0.140)	(0.176)	(0.146)	(0.217)	(0.220)
Ln(K/L)	0.024	0.035	0.100	0.144**	0.473***	0.381**
	(0.053)	(0.043)	(0.067)	(0.058)	(0.152)	(0.156)
Ln(Sales)	0.105***	0.081**	0.095**	0.093**	0.144**	0.177***
	(0.040)	(0.033)	(0.039)	(0.037)	(0.067)	(0.062)
Ln(Age)	-0.131	-0.067	-0.255***	-0.244***	-0.258***	-0.296***
	(0.083)	(0.067)	(0.095)	(0.075)	(0.094)	(0.096)
R&D stock, 10%	Benchmark	Benchmark	Benchmark	Benchmark	Benchmark	Benchmark
R&D stock, 20%	0.162	-0.096	0.119	-0.125	0.028	-0.259
	(0.145)	(0.090)	(0.161)	(0.132)	(0.187)	(0.220)
R&D stock, 30%	0.533***	0.174*	0.648***	0.332**	0.265	0.172
	(0.145)	(0.099)	(0.160)	(0.132)	(0.191)	(0.229)
R&D stock, 40%	0.648***	0.257**	0.578***	0.423***	0.207	0.323
	(0.154)	(0.113)	(0.168)	(0.137)	(0.199)	(0.237)
R&D stock, 50%	0.793***	0.347***	0.668***	0.621***	0.425**	0.567**
	(0.180)	(0.134)	(0.177)	(0.160)	(0.206)	(0.255)
R&D stock, 60%	0.784***	0.480***	0.887***	0.829***	0.436*	0.638**
	(0.182)	(0.140)	(0.184)	(0.160)	(0.229)	(0.260)
R&D stock, 70%	0.959***	0.635***	1.106***	0.998***	0.625***	0.805***
	(0.192)	(0.155)	(0.201)	(0.170)	(0.231)	(0.265)
R&D stock, 80%	1.259***	0.804***	1.379***	1.135***	0.941***	1.065***
	(0.206)	(0.175)	(0.210)	(0.185)	(0.234)	(0.275)
R&D stock, 90%	1.611***	1.285***	1.881***	1.610***	1.156***	1.438***
	(0.251)	(0.225)	(0.244)	(0.219)	(0.305)	(0.312)
R&D stock, 100%	1.858***	1.613***	2.260***	1.975***	0.966**	1.523***
	(0.316)	(0.302)	(0.296)	(0.284)	(0.383)	(0.376)
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS, NB and Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*), *R&D stock* dummy variables based on deciles of its distribution and other firm-level control variables. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Innovation and options volume – Internet bubble

Period	"1996 – 2000"			"2001 – 2004"		
Method	OLS	NB	Poisson	OLS	NB	Poisson
Dependent Var.	Ln(1+CITES)	CITES	CITES	Ln(1+CITES)	CITES	CITES
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.176***	0.134***	0.149***	0.161***	0.164***	0.179**
	(0.035)	(0.031)	(0.039)	(0.037)	(0.045)	(0.070)
InstOwn	0.002	0.089	-0.131	-0.303	-0.018	0.179
	(0.188)	(0.170)	(0.209)	(0.185)	(0.280)	(0.443)
Ln(K/L)	0.026	0.086	0.533***	0.048	0.125	0.769***
	(0.066)	(0.072)	(0.170)	(0.060)	(0.098)	(0.224)
Ln(Sales)	0.173***	0.144***	0.127	0.061	0.095	0.049
	(0.049)	(0.042)	(0.078)	(0.042)	(0.065)	(0.088)
Ln(Age)	-0.256***	-0.284***	-0.334***	0.121	-0.026	-0.291*
	(0.096)	(0.091)	(0.097)	(0.098)	(0.137)	(0.176)
Ln(R&D stock)	0.260***	0.291***	0.167*	0.243***	0.372***	0.205**
	(0.054)	(0.040)	(0.094)	(0.049)	(0.075)	(0.101)
Observations	1906	1906	1906	1365	1365	1365

Notes: This table presents estimates of OLS, NB and Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) on options volume (*Optvol*) and other firm-level control variables for the two subperiods 1996 – 2000 (during the "dot-com bubble") and 2001 – 2004 (after the "dot-com bubble"). Firms in columns: 501 in columns 1 – 3 and 398 in columns 4 – 6. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Innovation and options volume – Additional (financial) controls

Method	OLS		NB		Poisson	
Dependent Var.	Ln(1+CITES)	Ln(1+PATS)	CITES	PATS	CITES	PATS
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.110***	0.097***	0.118***	0.099***	0.152***	0.104**
	(0.022)	(0.019)	(0.023)	(0.020)	(0.044)	(0.051)
InstOwn	0.005	-0.290	-0.184	-0.212	-0.340	-0.226
	(0.242)	(0.183)	(0.299)	(0.210)	(0.298)	(0.292)
Ln(K/L)	-0.077	0.018	0.041	0.089	0.373***	0.255***
	(0.070)	(0.054)	(0.096)	(0.073)	(0.128)	(0.084)
Ln(Sales)	0.064	0.048	0.069	0.047	-0.177	0.045
	(0.060)	(0.045)	(0.059)	(0.053)	(0.117)	(0.088)
Ln(Age)	-0.041	0.048	-0.097	-0.064	-0.285**	-0.281**
	(0.113)	(0.086)	(0.122)	(0.091)	(0.131)	(0.124)
Ln(R&D stock)	0.274***	0.254***	0.269***	0.304***	0.492***	0.473***
	(0.069)	(0.054)	(0.061)	(0.059)	(0.118)	(0.085)
Illiquidity	-0.109*	-0.136**	-0.173**	-0.162***	-0.088	-0.032
	(0.062)	(0.052)	(0.068)	(0.055)	(0.086)	(0.067)
Leverage	0.548*	0.280	0.500	0.254	0.315	0.003
	(0.296)	(0.229)	(0.338)	(0.236)	(0.402)	(0.451)
Tobin's Q	-0.073	-0.047	-0.135	-0.085	-0.038	0.105
	(0.082)	(0.059)	(0.085)	(0.063)	(0.111)	(0.118)
ROA	-0.553	-0.531**	-0.871**	-0.758***	0.716	-0.205
	(0.373)	(0.251)	(0.361)	(0.286)	(0.449)	(0.448)
Capex	1.613	0.087	-0.344	-0.402	0.564	1.196
	(1.041)	(0.758)	(0.995)	(0.801)	(0.952)	(1.188)
Ln(1+Analyst coverage)	-0.003	0.044	0.029	0.050	0.070	0.039
	(0.080)	(0.064)	(0.083)	(0.064)	(0.073)	(0.058)
Observations	3271	3271	3271	3271	3271	3271

Notes: This table presents estimates of OLS, NB and Poisson panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*) and other (additional) firm-level control variables. *Illiquidity* is the natural logarithm of relative effective spread measured over firm i's fiscal year t, where relative effective spread is defined as the absolute value of the difference between the execution price and the midpoint of the prevailing bid-ask quote divided by the midpoint of the prevailing bid-ask quote; *Leverage* is the book value of debt (DLTT+DLC) divided by the book value of assets (AT); *Tobin's Q* is calculated as (market value of equity (PRCC_F × CSHO) plus book value of assets (AT) minus book value of equity (CEQ) minus balance sheet deferred taxes (TXDB)) divided by book value of assets (AT); *ROA* is the operating income before depreciation (OIDBP) divided by book value of assets (AT); *Capex* is defined as capital expenditures (CAPX) scaled by book value of assets (AT); and *Analyst coverage* is the arithmetic mean of the 12 monthly numbers of earnings forecasts for firm i extracted from the I/B/E/S summary file. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Innovation and options volume – External knowledge acquisition

Control variable	Collaboration frequency		Collaboration intensity		Acquisitions	
Dependent Var.	Ln(1+CITES)	Ln(1+PATS)	Ln(1+CITES)	Ln(1+PATS)	Ln(1+CITES)	Ln(1+PATS)
Method: OLS	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Optvol)	0.107***	0.090***	0.116***	0.158***	0.114***	0.156***
	(0.020)	(0.017)	(0.019)	(0.024)	(0.019)	(0.024)
InstOwn	-0.186	-0.252*	-0.164	-0.224	-0.153	-0.211
	(0.152)	(0.139)	(0.153)	(0.141)	(0.152)	(0.139)
Ln(K/L)	-0.012	0.040	-0.011	0.042	-0.014	0.040
	(0.048)	(0.043)	(0.049)	(0.044)	(0.048)	(0.044)
Ln(Sales)	0.129***	0.122***	0.121***	0.111***	0.130***	0.122***
	(0.036)	(0.033)	(0.037)	(0.034)	(0.036)	(0.033)
Ln(Age)	-0.062	-0.032	-0.059	-0.032	-0.057	-0.027
	(0.077)	(0.069)	(0.078)	(0.069)	(0.077)	(0.069)
Ln(R&D stock)	0.250***	0.193***	0.263***	0.210***	0.259***	0.204***
	(0.041)	(0.043)	(0.042)	(0.044)	(0.042)	(0.043)
Collaboration freq.	0.106***	0.129***				
	(0.035)	(0.032)				
Collaboration int.			-0.731*	-0.923***		
			(0.393)	(0.303)		
Acquisitions					-0.570**	-0.539**
					(0.290)	(0.257)
Observations	1446	1446	3271	3271	3271	3271

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*), *collaboration frequency*, *collaboration intensity*, *acquisitions* and other firm-level control variables. *Collaboration frequency* is the natural logarithm of (one plus) the number of R&D alliances formed over the previous five years (i.e., from $t-5$ to $t-1$); *Collaboration intensity* is the number of firms' jointly owned patents filed over the previous five years scaled by its total number of patents filed over the same period; *Acquisitions* is the acquisition expenditure (ACQ) divided by the book value of assets (AT). Firms in columns: 236 in columns 1 and 2; 548 in columns 3 – 8. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Innovation and options volume – High- versus low-tech industries

Dependent Var.	Ln(1+CITES)				Ln(1+PATS)			
Matched sample	All	All	High-tech	Low-tech	All	All	High-tech	Low-tech
Method: OLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Optvol)		0.186***				0.193***		
x Dummy for high-tech		(0.059)				(0.056)		
Ln(Optvol)	0.120***	0.004	0.144***	0.070	0.113***	-0.007	0.147***	0.046
	(0.039)	(0.052)	(0.049)	(0.058)	(0.034)	(0.049)	(0.040)	(0.051)
Dummy for high-tech	1.219**	0.932*			0.543	0.243		
	(0.485)	(0.510)			(0.578)	(0.605)		
InstOwn	-0.092	-0.053	-0.148	0.199	-0.164	-0.123	-0.287	0.208
	(0.189)	(0.188)	(0.248)	(0.252)	(0.170)	(0.170)	(0.231)	(0.216)
Ln(K/L)	-0.068	-0.071	0.085	-0.252*	-0.023	-0.025	0.090	-0.166
	(0.098)	(0.097)	(0.134)	(0.143)	(0.079)	(0.079)	(0.111)	(0.109)
Ln(Sales)	0.104*	0.091	0.180**	-0.076	0.100*	0.086*	0.163**	-0.031
	(0.062)	(0.062)	(0.087)	(0.081)	(0.053)	(0.052)	(0.075)	(0.065)
Ln(Age)	-0.138	-0.150*	-0.155	-0.156	-0.099	-0.111	-0.114	-0.119
	(0.085)	(0.084)	(0.113)	(0.118)	(0.073)	(0.072)	(0.097)	(0.102)
Ln(R&D stock)	0.167***	0.167***	0.212***	0.099**	0.143***	0.143***	0.158***	0.114***
	(0.034)	(0.034)	(0.050)	(0.043)	(0.028)	(0.027)	(0.041)	(0.035)
Observations	2906	2906	1453	1453	2906	2906	1453	1453

Notes: This table presents estimates of OLS panel regressions on a matched sample of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume *(Optvol)*, a dummy variable that equals one if a firm is operating in a high-tech industry (*Dummy for high-tech*), their interaction and other firm-level control variables. Firms in the matched sample: 547. Firms in columns 3 and 7: 311. Firms in column 4 and 8: 236. Robust standard errors are clustered by firm (in parentheses). The matched sample is constructed using nearest-neighbour matching with scores given by a probit model in which the dependent variable is *Dummy for high-tech*. The propensity score is estimated using the following firm characteristics: *Ln(Optvol)*, *InstOwn*, *Ln(K/L)*, *Ln(Sales)*, *Ln(Age)*, *Ln(R&D stock)*, *Illiquidity*, *Leverage*, *Tobin's Q*, *ROA*, *Capex*, *Ln(Analyst coverage)* and fixed effects. All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Innovation and options volume – Within-firm relationship

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Post	0.370**		0.277*	
	(0.146)		(0.118)	
Inclusion year −3		0.240		0.008
		(0.154)		(0.080)
Inclusion year −2		0.207		0.084
		(0.144)		(0.075)
Inclusion year −1		0.313**		0.135
		(0.142)		(0.098)
Inclusion year 1		0.243*		0.116
		(0.131)		(0.078)
Inclusion year 2		0.568***		0.271***
		(0.148)		(0.103)
Inclusion year 3		0.636***		0.335***
		(0.145)		(0.114)
Inclusion year 4		0.558***		0.526***
		(0.162)		(0.124)
InstOwn	-0.088	0.010	-0.159	-0.111
	(0.257)	(0.256)	(0.245)	(0.241)
Ln(K/L)	-0.058	-0.061	-0.011	-0.014
	(0.073)	(0.074)	(0.061)	(0.062)
Ln(Sales)	0.196***	0.198***	0.133***	0.134***
	(0.052)	(0.053)	(0.044)	(0.045)
Ln(Age)	-0.225**	-0.225**	-0.072	-0.070
	(0.112)	(0.112)	(0.087)	(0.087)
Ln(R&D stock)	0.211***	0.218***	0.252***	0.255***
	(0.043)	(0.044)	(0.048)	(0.048)
Observations	744	614	744	614

Notes: This table presents estimates of OLS panel regressions of within-firm changes in firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) before and after the option listing event. *Post* is a dummy variable equal to unity to indicate the post-listing period; *Inclusion year #* are dummy variables indicating the relative year around the listing event (the omitted category is the year of the event). Firms in columns: 93. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Innovation and options volume – Open interest as instrumental variable

Method	OLS (first stage)	2SLS (second stage)	
Dependent Var.	Ln(Optvol)	Ln(1+CITES)	Ln(1+PATS)
	(1)	(2)	(3)
Ln(Optvol) *(instr.)*		0.087***	0.102***
		(0.029)	(0.024)
InstOwn	1.185***	-0.028	-0.218
	(0.140)	(0.156)	(0.137)
Ln(K/L)	-0.205***	0.015	0.039
	(0.053)	(0.053)	(0.045)
Ln(Sales)	0.242***	0.156***	0.133***
	(0.028)	(0.041)	(0.033)
Ln(Age)	-0.412***	-0.121	-0.036
	(0.065)	(0.083)	(0.068)
Ln(R&D stock)	0.056**	0.273***	0.217***
	(0.028)	(0.046)	(0.044)
Ln(Open int.)	1.207***		
	(0.028)		
Observations	3271	3271	3271

Notes: This table presents estimates of 2SLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on options volume (*Optvol*) and other firm-level control variables, with the total open interest *Ln(Open int.)* as instrumental variable. Firms in all columns: 548. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The time period is 1996 – 2004 (with citations up to 2006). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Innovation and options volume – Managerial entrenchment (E-Index)

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Ln(Optvol)		-0.032**		-0.031**
x E-index		(0.015)		(0.013)
Ln(Optvol)	0.168***	0.160***	0.151***	0.144***
	(0.031)	(0.031)	(0.028)	(0.028)
E-index	-0.014	0.012	-0.016	0.009
(Entrenchment Index)	(0.034)	(0.034)	(0.029)	(0.029)
InstOwn	-0.020	-0.048	-0.015	-0.041
	(0.192)	(0.193)	(0.170)	(0.170)
Ln(K/L)	0.106	0.107	0.075	0.076
	(0.070)	(0.069)	(0.059)	(0.058)
Ln(Sales)	0.108**	0.103**	0.126***	0.120***
	(0.047)	(0.047)	(0.044)	(0.044)
Ln(Age)	-0.183**	-0.173*	-0.106	-0.098
	(0.090)	(0.090)	(0.084)	(0.084)
Ln(R&D stock)	0.126***	0.124***	0.116***	0.114***
	(0.028)	(0.028)	(0.024)	(0.024)
Observations	921	921	921	921

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*), managerial entrenchment (*E-Index*), their interaction and other firm-level control variables. Firms in columns: 331. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of three-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). The E-Index is an average of 6 provisions in the firm's charter (see Bebchuk, Cohen, and Ferrell, 2009). The measure is based on data from RiskMetrics in 1998, 2000, 2002 and 2004. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Tobin's Q and innovation

Dependent Var.: Tobin's Q Method: OLS	(1)	(2)
Ln(1+CITES)	0.050*	
	(0.026)	
Ln(1+PATS)		0.059*
		(0.033)
One-year lagged Tobin's Q	0.383***	0.383***
	(0.049)	(0.049)
Ln(Optvol)	0.236***	0.236***
	(0.037)	(0.037)
InstOwn	-0.052	-0.041
	(0.202)	(0.202)
Ln(K/L)	0.057	0.054
	(0.106)	(0.106)
Ln(Sales)	-0.221***	-0.225***
	(0.069)	(0.071)
Ln(Age)	-0.173	-0.175
	(0.112)	(0.111)
Ln(R&D stock)	0.030	0.028
	(0.039)	(0.039)
Leverage	-1.625***	-1.610***
	(0.297)	(0.295)
ROA	1.823***	1.830***
	(0.641)	(0.638)
Capex	0.598	0.634
	(2.470)	(2.476)
Observations	2658	2658

Notes: This table presents estimates of OLS regressions of firms' market value (*Tobin's Q*) on firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*), *one-year lagged Tobin's Q*, options volume (*Optvol*) and other firm-level control variables. Firms in all columns: 526. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies and time dummies. The time period is 1996–2004 (with citations up to 2006); * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Innovation and options volume – Controlling for possible mechanisms

Dependent Var.	Ln(1+CITES)		Ln(1+PATS)	
Method: OLS	(1)	(2)	(3)	(4)
Ln(Optvol)	0.216***	0.154***	0.192***	0.155***
	(0.050)	(0.050)	(0.042)	(0.043)
Competition		8.454***		5.384***
(1 − Lerner)		(2.619)		(1.764)
G-Index		-0.033		-0.032
(Governance Index)		(0.028)		(0.025)
Ln(CEO age)		-0.807**		-0.605*
		(0.396)		(0.325)
Δ ROA$_{t-1}$		-0.162		-0.123
		(0.406)		(0.307)
Ln(CEO vega)		0.038		0.113***
		(0.054)		(0.041)
Ln(CEO delta)		0.072		0.044
		(0.055)		(0.046)
InstOwn	-0.072	-0.092	-0.139	-0.145
	(0.293)	(0.296)	(0.259)	(0.266)
Ln(K/L)	0.023	0.055	0.079	0.099
	(0.102)	(0.101)	(0.084)	(0.080)
Ln(Sales)	0.231***	0.248***	0.227***	0.240***
	(0.082)	(0.080)	(0.072)	(0.071)
Ln(Age)	-0.147	-0.122	-0.102	-0.068
	(0.154)	(0.154)	(0.142)	(0.140)
Ln(R&D stock)	0.140**	0.142**	0.080	0.082
	(0.062)	(0.061)	(0.059)	(0.058)
Observations	1530	1530	1530	1530

Notes: This table presents estimates of OLS panel regressions of firms' patents weighted by the number of forward citations (*CITES*) and firms' unweighted patent counts (*PATS*) on product market competition (*Competition*), managerial entrenchment (*G-Index*), *CEO age*, lagged change in profitability (Δ ROA_{t-1}), stock-based compensation (*CEO vega* and *CEO delta*) and other firm-level control variables. Firms in columns: 285. Robust standard errors are clustered by firm (in parentheses). All regressions control for a full set of four-digit industry dummies, time dummies, and fixed effects by including pre-sample means of the dependent variable as proposed by Blundell, Griffith, and van Reenen (1999). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16: Descriptive statistics – Robustness tests

	Mean	StdDev	Min	Median	Max	Observations	Source
Co-patents/Patents$_{[t-5,t-1]}$	0.03	0.09	0	0	1	2391	USPTO
Leverage	0.16	0.16	0	0.13	0.91	3271	Compustat
Tobin's Q	2.97	2.75	0.40	2.13	39.12	3271	Compustat
Capex/Assets	0.05	0.05	0.00	0.04	0.53	3271	Compustat
Acquisition exp. (in $m)	85.1	458.3	-3557	0	8800.2	3271	Compustat
Average Open Interest	395.8	1056.3	0.03	92.4	13266.6	3271	OptionMetrics
Stock illiquidity	-5.52	2.14	-11.66	-5.45	2.89	3271	TAQ
Analyst Coverage	8.4	7.9	0	6.08	45.6	3271	I/B/E/S
R&D Alliances$_{[t-5,t-1]}$	3.7	15.2	0	0	270	1446	SDC Platinum
Entrenchment Index	2.1	1.1	0	2	5	921	RiskMetrics and Bebchuk et al. (2009)
CEO tenure	7.3	7.9	0	5	53	1845	ExecuComp
CEO cash comp. (in $000s)	1340.4	1491.9	0	961.9	43511.5	1845	ExecuComp

Appendix B

Appendix for "Modelling Electricity Swaps with Stochastic Forward Premium Models"

B.1 Spot Prices - One factor Model

We summarize the framework outlined in Lucia and Schwartz (2002). Spot electricity prices P_t are characterized as

$$P_t = f(t) + X_t \tag{B.1}$$

where $f(t)$ is a deterministic function, and X_t, is a mean-reverting stochastic process with constant volatility σ and, under the natural probability measure P follow:

$$dX_t = -kX_t dt + \sigma dZ_t^P \tag{B.2}$$

It can be shown (see Cartea and Figueroa (2005)) that under the risk-neutral probability measure Q it follows:

$$dX_t = k(\alpha^* - X_t)dt + \sigma dZ_t^Q \tag{B.3}$$

where dZ_t^Q are increments of standard independent Brownian motions Z_t^*, the mean reversion parameters are k and $X(0) = x_0$, where the drift terms are

$$\alpha^* = \frac{-\lambda\sigma}{k} \tag{B.4}$$

In this section we assume the Market Prices of Risk (MPR) of the electricity, which are λ respectively, to be constant over time. Under the risk-neutral measure the spot price P_t follows

$$P_t = f(t) + X_0 \exp(-kt) + \alpha^*(1 - \exp(-kt)) + \sigma \int_0^t \exp(k(s-t))dZ^Q \tag{B.5}$$

The distribution of P_t is Normal with mean given by:

$$E_0^Q(P_t) = f(t) + X_0 \exp(-kt) + \alpha^Q(1 - \exp(-kt)) \tag{B.6}$$

The value of any derivative security must be the expected value, under the risk-neutral measure, of its payoffs discounted to the valuation date at the risk-free rate. Assuming a constant risk-free rate r, the value at time zero of a forward contract on the spot price maturing at time T must be

$$V_0^T(P_T) = \exp(-rT)E_0^*[P_T - F_0(P_0, T)] \tag{B.7}$$

where $F_0(P_0, T)$ is the forward price set at time zero and T is the time to maturity. Since the value of a forward contract must be zero when it is first entered into, we obtain a closed form expression for computing forward prices with maturity T as follows

$$F_0(P_0, T) = E_0^*(P_T) = f(T) + (P_0 - f(0)) \exp(-kT) + \alpha^*(1 - \exp(-kT)) \tag{B.8}$$

The variance of the forward prices are given by

$$Var_0^T(P_T) = \frac{\sigma^2}{2k}(1 - \exp(-2kT)) \tag{B.9}$$

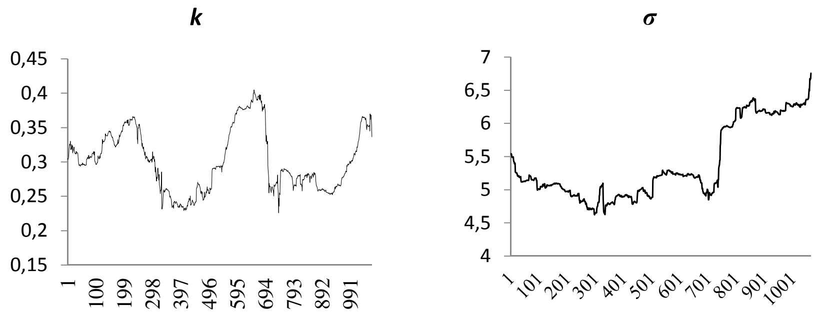
These results are for forward contracts providing electricity in a single point in time T. Given that the swap contract provides delivery of electricity during a period of time (e.g. during 31 days in January), we use (B.8) to generate prices during the full delivery period (e.g. we generate thirty-one forward prices in the cases of monthly contracts maturing in January and so on) , and we take the average. This average is the estimated swap price provided by the spot price model.

Table B.1: Estimation of the One Factor Model

	Coeff.	s.e.	t-stat
NEG	-79.86**	12.91	-6.18
@WEEKDAY=1	46.25**	0.77	60.09
@WEEKDAY=2	48.06**	0.64	75.30
@WEEKDAY=3	48.14**	0.67	71.49
@WEEKDAY=4	47.48**	0.66	71.76
@WEEKDAY=5	46.22**	0.71	65.28
@WEEKDAY=6	39.84**	0.60	65.99
@WEEKDAY=7	33.60**	0.68	49.21
$1-\hat{k}$	0.77**	0.03	29.70
$\hat{k}$	0.23**	0.025	8.79
$\hat{\sigma}$	5.86		
R^2	58.7		

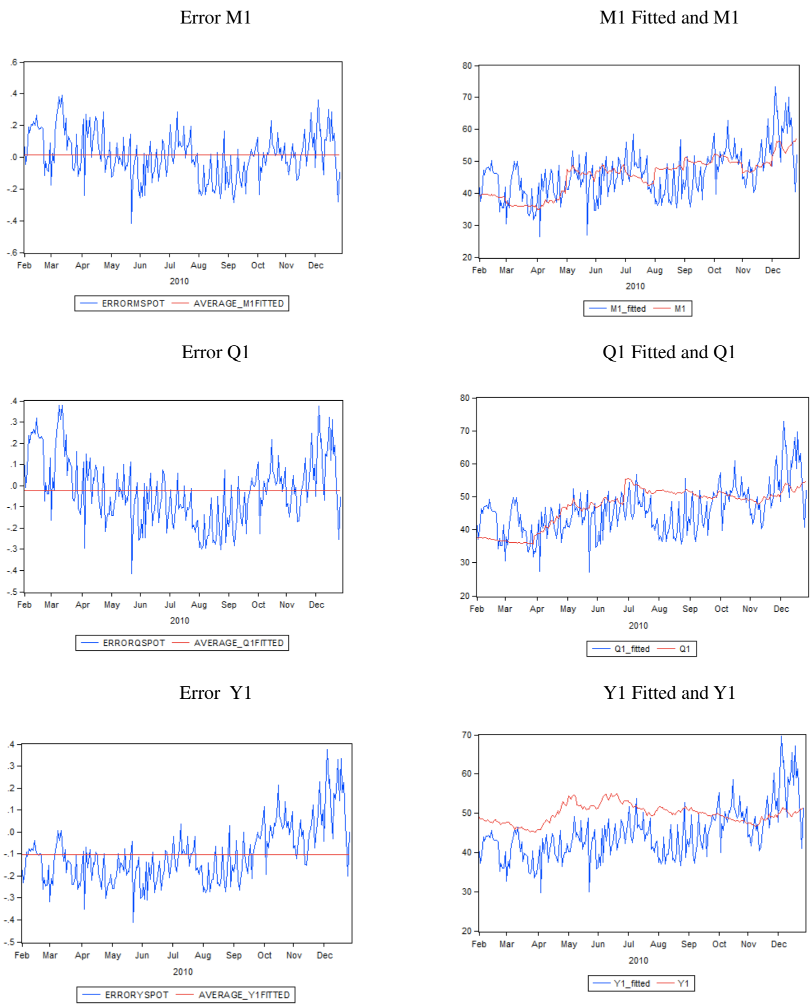
Notes: This table reports the results of regressions (B.1) and (B.2). The dependent variable is the average daily EEX spot price (EEX - Phelix Base Hr.01-24 E/Mwh). Our database spans from February 2, 2009 to December, 31 2012. Explanatory variables include day of the week dummies as well as the NEG variable which is a dummy variable taking into account negative electricity prices. It is equal to 1 if the price is negative (4 Oct 2009, 26 Dec 2009, 25 Dec 2012 y 26 Dec 2012) and it is zero otherwise. We estimate the coefficients by means of a regression robust to heteroskedasticity, and serial autocorrelation. The results presented correspond to the estimated coefficient, standard errors and t-statistics. The symbol * and ** denotes that the variable is significant at 5% and 1%, respectively.

Figure B.1: Parameter Stability



Notes: The figure depicts the recursive estimation of the parameters (mean reversion k and volatility σ) for the Spot Series 7 days EEX - Phelix Base Hr.01-24 E/Mwh. We use a 365-day rolling window. There are 1066 estimates. The last window is: 2/01/12-31/12/12. Average values are $k = 0.23$ and $\sigma = 5.86$.

Figure B.2: Swap Prices, Fitted Values and Errors in the Spot Model



Notes: Using the parameters estimated in Table B.1 we compute forward prices we obtain a closed form expression for computing forward prices with maturity T using Equation (B.8). We then generate swap theoretical prices (average of forward prices during the delivery period) using the previous equation for all contracts (denoted M1 fitted, Q1 fitted and Y1 fitted). As an illustration we present the results for $M+1$, $Q+1$ and $Y+1$ denoted M1, Q1 and Y1 (which are the most liquid contracts within each market segment) and compare them against market prices during 2010. The results for the other contracts and time periods are available on request.

B.2 HJM Model

We tested several different specifications for the HJM model. We present in this section the one based on the three more liquid contracts within each segment because in this case the model gives the best results . This model is a HJM-based multi-factor stochastic process for electricity swap prices under the real-world probability measure,

$$\frac{dF_i(t,T)}{F_i(t,T)} = \sum_{k=1}^{N} \alpha_{ki}(t,T) + \sigma_{ki}(t,T)dW_t^{ki} \tag{B.10}$$

Given that we work with the most liquid contract within each market segment we propose specific parameterizations for the volatility functions in (B.10) as follows

$$\frac{dF_i(t,T)}{F_i(t,T)} = \alpha_i + \sigma_{1i}(t,T_i)dW_t^{1i} \tag{B.11}$$

$$\sigma_{1i}(t,T_i) = \exp(-k_i(T_i - t)\sigma_{1i}) \tag{B.12}$$

where dW_t^{1i}, is an independent Brownian motion for all delivery periods, and $\sigma_{1i}(t,T_i)$ are volatility functions. We decide to use the parameterization (B.11)-(B.12) for the following reasons. The first question is how many factors are needed to realistically model the evolution of the forward curve. Extant empirical evidence (see Clewlow and Strickland (2000)) suggests that, in most cases, a minimum of two or three factors are needed to model the dynamics of the forward curve . So we decided to start with the most parsimonious representation, that is, one factor. Regarding the specific functions to be used, Equation (B.12) was chosen because of its analytical tractability and at the same time its ability in reflecting the well-known fact that short dated forward returns are more volatile than long dated forwards.

To calibrate this model we proceed as follows: We compute returns for our 3 contracts available in each case yearly contracts (Y1 to Y3), quarterly contracts (Q1 to Q3) and monthly contracts (M1 to M3). The volatility functions are recovered by eigenvector decomposition of the covariance matrix. The decomposition yields the set of independent factors that drive the evolution of the variables underlying the covariance matrix Σ. We decompose Σ into n (n=3) eigenvectors vi (size 3x1) and associated eigenvalues λ_i such that $\Sigma = RVR'$ where the columns of R are the eigenvectors and the principal diagonal in V contains the eigenvalues (other elements in V are zero). We only consider one eigenvalue. The first volatility function is computed by fitting the Equation (B.12) to the data $v_1\sqrt{\lambda_1}$.

The basic statistics of the return series are in Table B1. The average return in all cases is not statistically different from zero suggesting overall null drift for the swaps, and therefore we set $\alpha_i = 0$, $\forall i$. The estimated standard deviation is annualized by the number of trading days (250) and varies from 13% for the contract Y3 to 32% for the M1 contract. Volatility is usually higher for the closest to maturity contracts (Samuelson effect), confirming the well-known fact that short dated forward returns tend to be more volatile than long dated forwards. Figure B1 shows (sample period from 2004 to 2012) the term structure of the volatility for each market segment. The distribution of the returns has some skewness and deviates significantly from the normal distribution, as the very high kurtosis figures on Table B.1 suggest.

The eigenvalues resulting from the eigenvector decomposition tell us the importance of each eigenvector and hence the number of factors that we should include in our model. Thus the first eigenvector is the most important, explaining 87%, 89% and 94% of the total variation in the evolution of the swap curve for the monthly, quarterly and yearly contracts respectively.

Figure B.3 shows the first principal component function recovered from the above procedure for each contract type. The first principal component acts to shift the forward and also acts to tilt the swap curve. The most important factor (COMP1) is positive for all maturities, but decreasing with maturity. This implies that a positive shock to the system causes all prices to shift up but by decreasing amounts, depending on the maturity. The longer the maturity the smaller the increase in prices is. The Table B.3 presents the parameter estimates from the volatility function obtained in the Principal Component Analysis using equations for the entire sample 2004-2012.

Table B.3 presents in Panel A the LS estimates of the parameters from the volatility functions obtained in CPA using equation $\sigma_{1i}(t, T_i) = \exp(-k_i(T_i - t))\sigma_{1i}$ for the full sample period 2004-2012. Panel A reports the in-sample root mean squared pricing errors (RMSEs). Errors are computed daily based on the fitted swap prices based on estimated parameters in Panel A. We compute the volatility function implied by the HJM model and by the SFP model and compare the results against market prices. In the case of the HJM model, root-mean squared errors (RMSEs) are 6.05%, 19.32%,and 11.96% for monthly, quarterly and yearly contracts respectively. By contrast, the RMSEs for those contracts are 0.10%, 0.12% and 0.30% respectively in the case of the SFP model. The degree of fit of the SFP model is substantially higher than the HJM model. In the case of the first volatility function parameter σ_{1i} represents the overall volatility of the forward curve whilst parameter ki tells us how fast the forward volatility curve decreases with increasing maturity. The parameter σ_{1i} recaptures the annualized volatility averaged over all contracts of a given class. From Table B.1 it is easy to see that average volatility for annual, quarterly and monthly returns is 17%, 19% and 32% respectively. These are very close to estimated parameters σ_{1i} in Table B.1 which are 19%, 19% and 38% respectively. The reason of the proximity lies in the low values of the decay factor. Estimated values of this parameter k (0.15, 0.03 and 0.20) suggest a fairly slow decrease in volatility as time to maturity increases. Monthly prices present higher overall volatility and faster decrease in volatility with maturity, followed by quarterly and yearly prices. However, monthly prices present slower volatility attenuation than yearly prices. The degree of fit of the equation is high in the case of yearly prices (99%), followed by monthly (99%) and quarterly prices (83%). To check parameter stability we have repeated the calibration exercise using different subsamples 2006-2010 and 2010-2012. A comparison of the parameters can be observed in Figure B3, which shows the stability of the parameter estimates. Overall the results suggest that parameters are reasonably stable over time.

It is fair to say that, although the model seems to fit the volatility term structure to some extent, it is completely unable to recover the skewness and kurtosis observed in the empirical distributions. By contrast, the SFP model not only fits better the markets volatility term structure, but is also able to take into account the skew and kurtosis.

Table B.2: Descriptive Statistics of Returns

	M1	M2	M3	Q1	Q2	Q3	Y1	Y2	Y3
Mean	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Median	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	0.163	0.149	0.126	0.109	0.099	0.099	0.088	0.070	0.073
Minimum	-0.146	-0.163	-0.239	-0.062	-0.161	-0.074	-0.071	-0.063	-0.064
Std. Dev.	0.329	0.269	0.239	0.193	0.191	0.182	0.174	0.150	0.131
Skewness	0.175	-0.160	-1.837	0.326	-0.905	0.136	0.008	0.163	0.516
Kurtosis	9.999	14.152	37.444	9.801	24.640	10.586	9.424	10.023	14.248
p-val	0.080	0.133	0.195	0.511	0.985	0.984	0.962	0.595	0.998

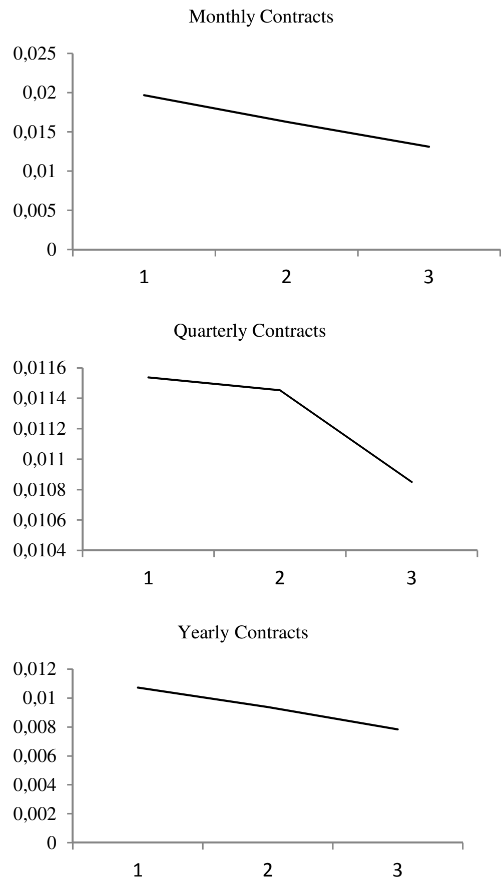
Notes: The table shows descriptive statistics of returns (1-day changes in the natural logs of swap prices) and the sample covariance matrix of these returns. We study three contracts for each market segment (yearly, quarterly and monthly), contracts M1 to M3, Q1 to Q3 and Y1 to Y3 from 6/1/2004 to 12/31/2012. The "Std. Dev." column reports the standard deviation of the series in annual terms. The nine series are corrected of the rolling effect by means of intervention analysis. p-val is the p-value for the test of zero mean.

Table B.3: Descriptive Statistics of Returns

	Yearly Contracts	Quarterly Contracts	Monthly Contracts
	Panel B: RMSEs		
HJM	6.05%	19.32%	11.96%
SFP	0.10%	0.12%	0.30%
	Panel A: Volatility Function		
σ_1	0.012	0.012	0.024
	(37.34)	(34.75)	(67.90)
K	-0.154	-0.030	-0.201
	(11.32)	(2.23)	(26.09)
R^2	0.992	0.834	0.998

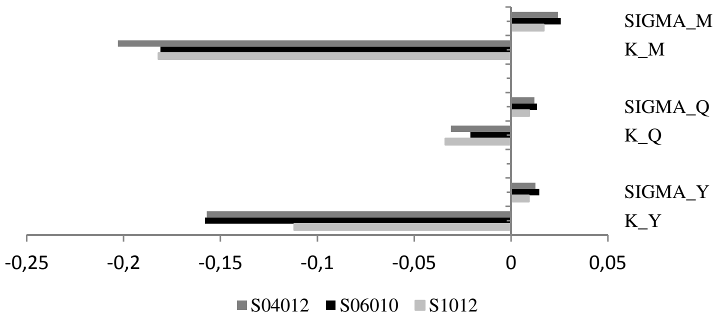
Notes: The table presents in Panel A the LS estimates of the parameters from the volatility functions obtained in CPA using equation $\sigma_{1i}(t, T_i) = \exp(-k_i(T_i - t))\sigma_{1i}$ for the full sample period 2004-2012. t-statistics are presented in parenthesis. Panel A reports in-sample root mean squared pricing errors for HJM and SFP models. In the case of HJM, the errors are computed daily based on the fitted prices of swap based on estimated parameters in Panel B.

Figure B.3: Volatility Functions



Notes: The Figure shows the first principal component functions recovered from the above procedure for each contract type (M1, M2, M3; Q1, Q2, Q3 and Y1,Y2, Y3). Sample period 2004-2012.

Figure B.4: Stability of the parameters by subsamples



	S04012	S06010	S1012
SIGMA_M	0.0242	0.0257	0.0169
K_M	-0.203	-0.181	-0.182
SIGMA_Q	0.012	0.0134	0.0094
K_Q	-0.031	-0.021	-0.034
SIGMA_Y	0.0126	0.0146	0.0093
K_Y	-0.157	-0.158	-0.112

Notes: To check parameter stability we have repeated the calibration exercise using different subsamples 2006-2010 and 2010-2012. The plot presents a comparison of the parameters.

B.3 Value at Risk

In order to compare the performance of alternative models in comparison with the SFP model, we compute Value-at-Risk (VaR) at different probability levels over a one-day horizon. Given a portfolio P, a time T and a probability level Q, a loss L^* is selected, at which exists a probability Q that effective losses L, are at most L^* in period T. The loss L^* is portfolios VaR. Formally,

$$Prob[L^* \geqslant L] = Q \tag{B.13}$$

and therefore VaR_Q is a quintile of assets returns probability density function, which defines the maximum expected loss with confidence level Q. In the following, and to be consistent with the empirical evidence in our sample, we assume that the expected one-day swap return is zero. A comparison of the VaR for standardized returns and for different probability levels, Q is shown in Figure B.5, for the Normal distribution and for the NIG distribution with different kurtosis parameter values. In the case of relatively low significance levels (90% and 95%), the values of VaR_Q^{Normal} tend to be higher (in absolute terms) than those of VaR_Q^{NIG}, so the latter measure will probably underestimate risk. However, in the case of high significance levels (99% and beyond) there is a very substantial difference between the two measures, because the VaR_Q^{Normal}

strongly underestimates the risk in comparison with VaR_Q^{NIG}. The difference between the two measures, for a given Q, is higher; the closer to the unity is the kurtosis parameter ξ.

For the computation of the 1-day VaR for each swap contract, we proceed as follows. We assume that the innovations in the spot model and in the HJM model are normal. However, and given the limited success of the spot price model in our sample, we employ errors from the HJM model for the VaR calculations. Therefore, we compute $VaR_Q^{Normal}(i,T)$ as follows

$$VaR_Q^{Normal}(i,T) = k(\sigma_{i,T}\sqrt{\Delta t}) \tag{B.14}$$

where the factor k depends on Q as presented in Figure 3.4 and $\sigma_{i,T}$ is the volatility of the innovations of the forward prices generated by means of the HJM model.

To compute the VaR in the case of the SFP model, we consider that each swap can be thought as a portfolio containing two stochastic factors and therefore its VaR should be computed using the standard VaR formula for a portfolio (Jorion, 2007). Using Equation (3.1), we define the VaR for the swap in market segment i and maturity T as follows

$$VaR_Q^{NIG}[F_i(t,T)] = \sqrt{VaR_Q[\overline{F}_i(t)]^2 + VaR_Q[\gamma_i(t,T)]^2 + 2Cov(\overline{F}_i(t),\gamma_i(t,T))} \tag{B.15}$$

We compute the cumulative distributions functions for the NIG processes driving $\overline{F}_i(t)$ and $\gamma_i(t,T))$ by means of numerical simulation. We compute the VaR for each component as follows

$$VaR_Q[\overline{F}_i(t)] = k_{\xi,\chi}(\theta_{\overline{F}_i}\sqrt{\Delta t}) \tag{B.16}$$

$$VaR_Q[\gamma_i(t,T)] = k_{\xi,\chi}(\theta_{\gamma_i(t,T)}\sqrt{\Delta t}) \tag{B.17}$$

where the factor $k_{\xi,\chi}$ depends on Q and on the skewness and kurtosis parameters. The volatilities are the residual standard errors obtained from Equation (3.12) and reported in Table 3.5. We obtain the covariance term from matrix (see Table 3.6) and it is computed as follows

$$Cov(\overline{F}_i(t),\gamma_i(t,T)) = \theta_{\overline{F}_i} \times \theta_{\gamma_i(T)} \times \omega_{\overline{F}_i,\gamma_i(T)} \tag{B.18}$$

Next, we compare the Failure Ratios (FR) of these two alternative models in our sample. We define FR as

$$FR = \frac{N/T}{c}$$

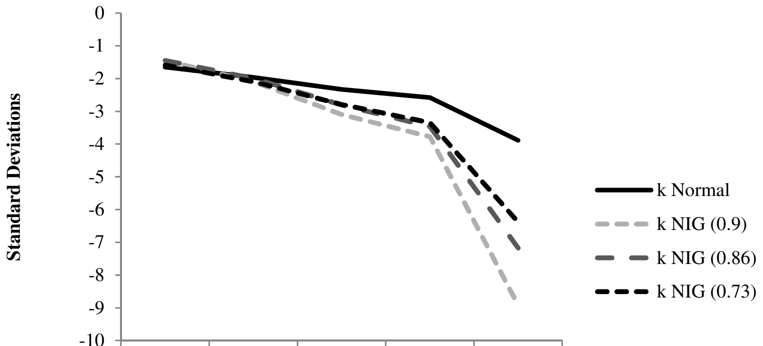
where $1-c$ is the confidence level, T is number of time periods (e.g. $T = 100$ days) and a failure appears when the realized loss (negative return) is larger than the VaR forecast. If the model producing the VaR forecasts is right in assessing the risk, we expect that $FR = 1$. On the other hand if the model tends to underestimate, (overestimate) risk, then $FR > 1$ ($FR < 1$). To formally test the statistical difference from 1 of the estimated FR , we use a variation of Kupiec (1995) test, suggested by Campbell (2006). Under the assumption that the VaR under consideration is accurate, the z-statistic has an approximate standard normal distribution and has a known exact finite sample distribution. The z statistic is actually the Wald variant of the likelihood ratio statistic proposed by Kupiec (1995). One potential advantage of the Wald test over the likelihood ratio test is that it is well-dened in the case that no VaR violations occur. Kupiecs test is not dened in this case. Moreover, the possibility that no violations occur in a relatively short period, is not trivial. The z-statistic is

$$Z = \frac{\sqrt{T}\left(\dfrac{N}{T} - c\right)}{\sqrt{c(1-c)}} \tag{B.19}$$

A positive (negative) z statistic indicates that the model tends underestimate (over-estimate) risk. We present the results in Table 3.9. It may be seen that VaR_Q^{Normal}, calculated under the assumption of normality tends to understate risk, and this under-statement is very strong for high confidence levels (99.5% and 99.99%) suggesting that tail risk is severely underestimated. On the other hand VaR_Q^{NIG} tends to mildly overestimate risk at relatively low significance levels but it is able to properly account for extreme tail risk. It is worth noting that the overestimation of risk provided by the NIG distribution is proportionally much lower than the underestimation of risk produced by the normal distribution.

One important practical implication of our results is as follows. We know that VaR models allow users to control risk and decide how to allocate limited resources. Financial intermediaries impose a capital charge to traders based on risk-adjusted capital. This creates a natural incentive for traders to take a position only when they have strong views on markets. If they have no views, they should abstain from trading. Our results suggest that the risk adjusted capital for traders using EEX swap electricity contracts should be adjusted upwards in comparison with the standard practice based on the normality assumption. Traders should also rationally adjust positions as risk changes (in the face of an increasingly volatile environment a sensible response is to scale down positions). Furthermore and given that VaR is also a performance evaluation tool, the evaluators of the performance of the traders should adjust their measures accordingly.

Figure B.5: Critical values



	0,95	0,98	0,99	0,995	0,999
k Normal	-1,65	-1,96	-2,33	-2,58	-3,89
k NIG (0.9)	-1,45	-2,09	-3,09	-3,77	-8,94
k NIG (0.86)	-1,44	-1,99	-2,8	-3,46	-7,18
k NIG (0.73)	-1,58	-2,09	-2,79	-3,34	-6,39

Notes: The figure plots the critical values of Value-at-Risk for Different Levels of Significance (Q): Standardized Normal and NIG distributions

Bibliography

Acs, Z. J., L. Anselin, and A. Varga (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy 31*(7), 1069–1085.

Acs, Z. J. and D. B. Audretsch (1988). Innovation in large and small firms: An empirical analysis. *American Economic Review 78*(4), 678–690.

Admati, A. R. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies 1*(1), 3–40.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005). Competition and innovation: an inverted-u relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Aghion, P., J. Van Reenen, and L. Zingales (2013). Innovation and institutional ownership. *American Economic Review 103*(1), 277–304.

Albert, M., D. Avery, F. Narin, and P. McAllister (1991). Direct validation of citation counts as indicators of industrially important patents. *Research Policy 20*(3), 251 – 259.

Alexander, C. (1999). Correlation and cointegration in energy markets. *Managing energy price risk 2*.

Amin, K., V. Ng, and S. Pirrong (1995). Vvaluing energy derivatives, v chapter 3 in managing energy price risk, 57'70.

Amore, M. D., C. Schneider, and A. Žaldokas (2013). Credit supply and corporate innovation. *Journal of Financial Economics 109*(3), 835–855.

Andresen, A., S. Koekebakker, and S. Westgaard (2010a). Modeling electricity forward prices using the multivariate normal inverse gaussian distribution. *J. Energy Mark 3*(3), 3–25.

Andresen, A., S. Koekebakker, and S. Westgaard (2010b). Modeling electricity forward prices using the multivariate normal inverse gaussian distribution. *J. Energy Mark 3*(3), 3–25.

Anton, J. J. and D. A. Yao (2002). The sale of ideas: Strategic disclosure, property rights, and contracting. *The Review of Economic Studies 69*(3), 513–531.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical finance 9*(3), 203–228.

Asker, J., J. Farre-Mensa, and A. Ljungqvist (2015). Corporate investment and stock market listing: A puzzle? *Review of Financial Studies 28*(2), 342–390.

Audet, N., P. Heiskanen, J. Keppo, and I. Vehviläinen (2002). Modelling of electricity forward curve dynamics. *Modelling Prices in Competitive Electricity Markets, DW Bunn, ed, John Wiley & Sons*, 251–265.

Aumann, R. J. and R. Serrano (2008). An economic index of riskiness. *Journal of Political Economy 116*(5), 810–836.

Balbás, A., B. Balbás, and R. Balbás (2010). Minimizing measures of risk by saddle point conditions. *Journal of computational and applied mathematics 234*(10), 2924–2931.

Balbás, A., B. Balbás, and R. Balbás (2013). Good deals in markets with friction. *Quantitative Finance 13*(6), 827–836.

Balbás, A., R. Balbás, and J. Garrido (2010). Extending pricing rules with general risk functions. *European Journal of Operational Research 201*(1), 23–33.

Barndorff-Nielsen, O. E. and K. Prause (2001). Apparent scaling. *Finance and Stochastics 5*(1), 103–113.

Bebchuk, L., A. Cohen, and A. Ferrell (2009). What matters in corporate governance? *Review of Financial Studies 22*(2), 783–827.

Benth, F. E., J. S. Benth, and S. Koekebakker (2008). *Stochastic modelling of electricity and related markets*, Volume 11. World Scientific.

Benth, F. E., Á. Cartea, and R. Kiesel (2008). Pricing forward contracts in power markets by the certainty equivalence principle: explaining the sign of the market risk premium. *Journal of Banking & Finance 32*(10), 2006–2021.

Benth, F. E. and S. Koekebakker (2008). Stochastic modeling of financial electricity contracts. *Energy Economics 30*(3), 1116–1157.

Bernstein, S. (2015). Does going public affect innovation? *The Journal of Finance 70*(4), 1365–1403.

Bertrand, M. and S. Mullainathan (2003). Enjoying the quiet life? corporate governance and managerial preferences. *Journal of Political Economy 111*(5), 1043–1075.

Bessembinder, H. and M. L. Lemmon (2002). Equilibrium pricing and optimal hedging in electricity forward markets. *the Journal of Finance 57*(3), 1347–1382.

Bessen, J. (2009). Tool for matching assignee names. National Bureau of Economic Research PDP Project.

Bhattacharya, S. and J. R. Ritter (1983). Innovation and communication: Signalling with partial disclosure. *The Review of Economic Studies 50*(2), 331–346.

Bjerksund, P., H. Rasmussen, and G. Stensland (2010). Valuation and risk management in the norwegian electricity market. In *Energy, natural resources and environmental economics*, pp. 167–185. Springer.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal 31*(4), 36–72.

Blundell, R., R. Griffith, and J. van Reenen (1999). Market share, market value and innovation in a panel of british manufacturing firms. *The Review of Economic Studies 66*(3), 529–554.

Blundell, R. W. and J. L. Powell (2004). Endogeneity in semiparametric binary response models. *The Review of Economic Studies 71*(3), 655–679.

Borovkova, S. and H. Geman (2006a). Analysis and modelling of electricity futures prices. *Studies in Nonlinear Dynamics & Econometrics 10*(3).

Borovkova, S. and H. Geman (2006b). Seasonal and stochastic effects in commodity forward curves. *Review of Derivatives Research 9*(2), 167–186.

Brace, A., M. Musiela, et al. (1997). The market model of interest rate dynamics. *Mathematical finance 7*(2), 127–155.

Bushee, B. J. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review 73*(3), 305–333.

Campbell, S. D. (2006). A review of backtesting and backtesting procedures. *The Journal of Risk 9*(2), 1.

Cao, H. H. (1999). The effect of derivative assets on information acquisition and price behavior in a rational expectations equilibrium. *Review of Financial Studies 12*(1), 131–163.

Capasso, M., L. Alessi, M. Barigozzi, and G. Fagiolo (2009). On approximating the distributions of goodness-of-fit test statistics based on the empirical distribution function: The case of unknown parameters. *Advances in Complex Systems 12*(02), 157–167.

Cartea, A. and M. G. Figueroa (2005). Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance 12*(4), 313–335.

Casassus, J. and P. Collin-Dufresne (2005). Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance 60*(5), 2283–2331.

Cassiman, B. and R. Veugelers (2006). In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. *Management Science 52*(1), 68–82.

Chakravarty, S., H. Gulen, and S. Mayhew (2004). Informed trading in stock and option markets. *The Journal of Finance 59*(3), 1235–1257.

Chavez-Demoulin, V., P. Embrechts, and J. Nešlehová (2006). Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking & Finance 30*(10), 2635–2658.

Chen, Q., I. Goldstein, and W. Jiang (2007). Price informativeness and investment sensitivity to stock price. *Review of Financial Studies 20*(3), 619–650.

Chowdhry, B. and V. Nanda (1991). Multimarket trading and market liquidity. *Review of Financial Studies 4*(3), 483–511.

Clewlow, L. and C. Strickland (2000). *Energy derivatives: pricing and risk management.* Lacima Publ.

Coad, A. and R. Rao (2008). Innovation and firm growth in high-tech sectors: A quantile regression approach. *Research Policy 37*(4), 633–648.

Cohen, W. M. and R. C. Levin (1989). Empirical studies of innovation and market structure. In R. Schmalensee and R. D. Willig (Eds.), *Handbook of Industrial Organization* (2nd ed.). North-Holland.

Cohen, W. M., R. R. Nelson, and J. P. Walsh (2000). Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). National Bureau of Economic Research Working Paper 7552.

Coles, J. L., N. D. Daniel, and L. Naveen (2006). Managerial incentives and risk-taking. *Journal of Financial Economics 79*(2), 431–468.

Core, J. and W. Guay (2002). Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting Research 40*(3), 613–630.

Cortazar, G. and E. S. Schwartz (1994). The valuation of commodity contingent claims. *The Journal of Derivatives 1*(4), 27–39.

Datar, S. and C. Reavis (2003). Novartis pharma: The business unit model. Harvard Business School Case Study 9-101-030.

Delbaen, F. (2009). Risk measures for non-integrable random variables. *Mathematical Finance 19*(2), 329–333.

Di Poto, G. and E. Fanone (2012). Estimating a lévy multifactor market model for electricity futures markets by using independent component analysis. *The Journal of Energy Markets 5*(4), 33.

Dow, J. and G. Gorton (1997). Stock market efficiency and economic efficiency: Is there a connection? *The Journal of Finance 52*(3), 1087–1129.

Durnev, A., R. Morck, and B. Yeung (2004). Value-enhancing capital budgeting and firm-specific stock return variation. *The Journal of Finance 59*(1), 65–105.

Easley, D., M. O'Hara, and P. Srinivas (1998). Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance 53*(2), 431–465.

Edmans, A., X. Gabaix, and A. Landier (2009). A multiplicative model of optimal ceo incentives in market equilibrium. *Review of Financial Studies 22*(12), 4881–4917.

Edmans, A. and G. Manso (2011). Governance through trading and intervention: A theory of multiple blockholders. *Review of Financial Studies 24*(7), 2395–2428.

Embrechts, P., R. Frey, and A. McNeil (2005). Quantitative risk management. *Princeton*

Series in Finance, Princeton 10.

Embrechts, P., G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj (2014). An academic response to basel 3.5. *Risks 2*(1), 25–48.

Escribano, A., J. Ignacio Peña, and P. Villaplana (2011). Modelling electricity prices: International evidence*. *Oxford bulletin of economics and statistics 73*(5), 622–650.

Fang, V. W., X. Tian, and S. Tice (2014). Does stock liquidity enhance or impede firm innovation? *The Journal of Finance 69*(5), 2085–2125.

Faure-Grimaud, A. and D. Gromb (2004). Public trading and private incentives. *Review of Financial Studies 17*(4), 985–1014.

Ferreira, D., M. A. Ferreira, and C. C. Raposo (2011). Board structure and price informativeness. *Journal of Financial Economics 99*(3), 523–545.

Fleten, S.-E. and J. Lemming (2003). Constructing forward price curves in electricity markets. *Energy Economics 25*(5), 409–424.

Föllmer, H. and A. Schied (2002). Convex measures of risk and trading constraints. *Finance and stochastics 6*(4), 429–447.

Forsberg, L. and T. Bollerslev (2002). Bridging the gap between the distribution of realized (ecu) volatility and arch modelling (of the euro): the garch-nig model. *Journal of Applied Econometrics 17*(5), 535–548.

Foucault, T. and T. Gehrig (2008). Stock price informativeness, cross-listings, and investment decisions. *Journal of Financial Economics 88*(1), 146–168.

Francis, B. B., I. Hasan, and Z. Sharma (2011). Incentives and innovation: Evidence from CEO compensation contracts. Bank of Finland Research Discussion Paper 17.

Freeman, C. and L. Soete (1997). *The Economics of Industrial Innovation* (3rd ed.). MIT Press.

Frestad, D. (2008). Common and unique factors influencing daily swap returns in the nordic electricity market, 1997–2005. *Energy Economics 30*(3), 1081–1097.

Frestad, D., F. E. Benth, and S. Koekebakker (2010). Modeling term structure dynamics in the nordic electricity swap market. *The Energy Journal*, 53–86.

Geman, H. and A. Roncoroni (2006). Understanding the fine structure of electricity prices*. *The Journal of Business 79*(3), 1225–1261.

Gibbons, R. and K. J. Murphy (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of Political Economy 100*(3), 468–505.

Gillan, S. L., J. C. Hartzell, and R. Parrino (2009). Explicit versus implicit contracts: Evidence from CEO employment agreements. *The Journal of Finance 64*(4), 1629–1655.

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Gompers, P., J. Ishii, and A. Metrick (2003). Corporate governance and equity prices. *The Quarterly Journal of Economics 118*(1), 107–156.

González-Pedraz, C., M. Moreno, and J. I. Peña (2014). Tail risk in energy portfolios. *Energy Economics 46*, 422–434.

Graham, J. R., C. R. Harvey, and S. Rajgopal (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics 40*(1-3), 3–73.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. National Bureau of Economic Research Working Paper 3301.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2001). The NBER patent citation data file: Lessons, insights and methodological tools. National Bureau of Economic Research Working Paper 8498.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics 36*(1), 16–38.

Hall, B. H. and J. Lerner (2010). The financing of R&D and innovation. In B. H. Hall and N. Rosenberg (Eds.), *Handbook of The Economics of Innovation* (1st ed.). North-Holland.

Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel (1999). Citation frequency and the value of patented inventions. *The Review of Economics and Statistics 81*(3), 511–515.

Hart, O. D. (1983). The market mechanism as an incentive scheme. *The Bell Journal of Economics 14*(2), 366–382.

Hausman, J. A., B. H. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica 52*(4), 909–938.

He, J. J. and X. Tian (2013). The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics 109*(3), 856–878.

Heath, D., R. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, 77–105.

Hilliard, J. E. and J. Reis (1998). Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot. *Journal of Financial and Quantitative Analysis 33*(01), 61–86.

Hirshleifer, D., A. Low, and S. H. Teoh (2012). Are overconfident CEOs better innovators? *The Journal of Finance 67*(4), 1457–1498.

Holmström, B. (1989). Agency costs and innovation. *Journal of Economic Behavior & Organization 12*(3), 305–327.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies 66*(1), 169–182.

Holmström, B. and J. Tirole (1993). Market liquidity and performance monitoring. *Journal of Political Economy 101*(4), 678–709.

Hsu, P.-H., X. Tian, and Y. Xu (2014). Financial development and innovation: Cross-country evidence. *Journal of Financial Economics 112*(1), 116–135.

Hu, J. (2014). Does option trading convey stock price information? *Journal of Financial Economics 111*(3), 625–645.

Jaffe, A. B. and M. Trajtenberg (2002). *Patents, citations, and innovations: A window on the knowledge economy.* MIT Press.

Jenter, D. and K. Lewellen (2014). Performance-induced CEO turnover. Stanford University Working Paper 3054.

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*, Volume 3. McGraw-Hill New York.

Karlis, D. (2002). An em type algorithm for maximum likelihood estimation of the normal–inverse gaussian distribution. *Statistics & Probability Letters 57*(1), 43–52.

Katila, R. and S. Shane (2005). When does lack of resources make new firms innovative? *Academy of Management Journal 48*(5), 814–829.

Khanna, N., S. Slezak, and M. Bradley (1994). Insider trading, outside search, and resource allocation: Why firms and society may disagree on insider trading restrictions. *Review of Financial Studies 7*(3), 575–608.

Kiesel, R., G. Schindlmayr, and R. H. Börger (2009). A two-factor model for the electricity forward market. *Quantitative Finance 9*(3), 279–287.

Koekebakker, S. and F. Ollmar (2005). Forward curve dynamics in the nordic electricity market. *Managerial Finance 31*(6), 73–94.

Konno, H. and R. Yamamoto (2005). Global optimization versus integer programming in portfolio optimization under nonconvex transaction costs. *Journal of Global Optimization 32*(2), 207–219.

Kothari, S. (2001). Capital markets research in accounting. *Journal of Accounting and Economics 31*(1-3), 105–231.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives 3*(2).

Kupper, M. and G. Svindland (2011). Dual representation of monotone convex functions on l0. *Proceedings of the American Mathematical Society 139*(11), 4073–4086.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

Lerner, J., M. Sorensen, and P. Strömberg (2011). Private equity and long-run investment: The case of innovation. *The Journal of Finance 66*(2), 445–477.

Lerner, J. and J. Wulf (2007). Innovation and incentives: Evidence from corporate R&D.

The Review of Economics and Statistics 89(4), 634–644.

Levin, R. C., A. K. Klevorick, R. R. Nelson, S. G. Winter, R. Gilbert, and Z. Griliches (1987). Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity 1987*(3), 783–831.

Lillestol, J. (2000). Risk analysis and the nig distribution. *Journal of Risk 2*, 41–56.

Ljungqvist, A. and W. J. Wilhelm (2003). IPO pricing in the dot-com bubble. *The Journal of Finance 58*(2), 723–752.

Lucia, J. J. and E. S. Schwartz (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of derivatives research 5*(1), 5–50.

Luo, Y. (2005). Do insiders learn from outsiders? evidence from mergers and acquisitions. *The Journal of Finance 60*(4), 1951–1982.

Mansfield, E. (1986). Patents and innovation: An empirical study. *Management Science 32*(2), 173–181.

Mansini, R., W. Ogryczak, and M. G. Speranza (2007). Conditional value at risk and related linear programming models for portfolio optimization. *Annals of operations research 152*(1), 227–256.

Mayhew, S. and V. Mihov (2004). How do exchanges select stocks for option listing? *The Journal of Finance 59*(1), 447–471.

Miltersen, K. R. (2003). Commodity price modelling that matches current observables: A new approach.

Miltersen, K. R. and E. S. Schwartz (1998). Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis 33*(01), 33–59.

Narayanan, M. P. (1985). Managerial incentives for short-term results. *The Journal of Finance 40*(5), 1469–1484.

Ni, S. X., J. Pan, and A. M. Poteshman (2008). Volatility information trading in the option market. *The Journal of Finance 63*(3), 1059–1091.

OECD (2011). ISIC Rev. 3 Technology Intensity Definition, Classification of manufacturing industries into categories based on R&D intensities. OECD Publishing.

OECD (2013). OECD Science, Technology and Industry Scoreboard. OECD Publishing.

Ogryczak, W. and A. Ruszczynski (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization 13*(1), 60–78.

Øigård, T. A., A. Hanssen, R. E. Hansen, and F. Godtliebsen (2005). Em-estimation and modeling of heavy-tailed processes with the multivariate normal inverse gaussian distribution. *Signal Processing 85*(8), 1655–1673.

Pakes, A. and Z. Griliches (1980). Patents and R&D at the firm level: A first report. *Economics Letters 5*(4), 377–381.

Pan, J. and A. M. Poteshman (2006). The information in option volume for future stock prices. *Review of Financial Studies 19*(3), 871–908.

Porter, M. E. (1992). Capital disadvantage: America's failing capital investment system. *Harvard business review 70*(5), 65–82.

Quinn, J. A., J. D. Reitzes, and A. C. Schumacher (2005). Forward and spot prices in electricity and gas markets. In *Obtaining the best from Regulation and Competition*, pp. 109–134. Springer.

Rockafellar, R. T., S. Uryasev, and M. Zabarankin (2006). Generalized deviations in risk analysis. *Finance and Stochastics 10*(1), 51–74.

Roll, R., E. Schwartz, and A. Subrahmanyam (2009). Options trading activity and firm valuation. *Journal of Financial Economics 94*(3), 345–360.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rosenbaum, P. R. and D. B. Rubin (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association 79*(387), 516–524.

Ross, S. A. (1976). Options and efficiency. *The Quarterly Journal of Economics 90*(1), 75–89.

Rothaermel, F. T. and M. T. Alexandre (2009). Ambidexterity in technology sourcing: The moderating role of absorptive capacity. *Organization Science 20*(4), 759–780.

Rudin, W. (1973). Functional analysis. 1973.

Rydberg, T. H. (1999). Generalized hyperbolic diffusion processes with applications in finance. *Mathematical Finance 9*(2), 183–201.

Schmidt, K. M. (1997). Managerial incentives and product market competition. *The Review of Economic Studies 64*(2), 191–213.

Schwartz, E. and J. E. Smith (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science 46*(7), 893–911.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance 52*(3), 923–973.

Shleifer, A. and R. W. Vishny (1990). Equilibrium short horizons of investors and firms. *American Economic Review 80*(2), 148–153.

Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of Political Economy 96*(1), 61–80.

Stein, J. C. (1989). Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *The Quarterly Journal of Economics 104*(4), 655–669.

Subrahmanyam, A. and S. Titman (1999). The going-public decision and the development of financial markets. *The Journal of Finance 54*(3), 1045–1082.

Supatgiat, C., R. Q. Zhang, and J. R. Birge (2001). Equilibrium values in a competitive power exchange market. *Computational Economics 17*(1), 93–121.

Tian, X. and T. Y. Wang (2014). Tolerance for failure and corporate innovation. *Review of Financial Studies 27*(1), 211–255.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The RAND Journal of Economics 21*(1), 172–187.

Trolle, A. B. and E. S. Schwartz (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *Review of Financial Studies 22*(11), 4423–4461.

Urzua, C. (1997). Omnibus tests for multivariate normality based on a class of maximum entropy distributions! *Advances in Econometrics 12*, 341–353.